Date: March 1st., 1986

To: Dr. Harvey Gonick

From: Roderick J.A. Little, Ph.D.
SBCC Consulting Clinic
Biomathematics Department, UCLA.

Subject: Statistical Review of Blood Lead - Blood Pressure studies.

Here are my impressions of the draft EPA Addendum "Lead Effects on Cardiovascular Function and Stature" and associated submissions. Unfortunately your time constraints do not allow an exhaustive critique or further analysis of the data.

1. Main Conclusions.

1.1. Important discrepancies pointed out by Ware and Royall between the Schwartz analyses in his June 26 1985 memorandum to David Weil and the Dupont analysis of May 9, 1985 remain to be explained. The later Schwartz analysis of October 2nd, 1985 does not replicate the Dupont analysis using exactly the same variables and data base, and Dupont have not attempted to replicate the Schwartz analysis. These analyses yield markedly different blood lead coefficients, with or without controls for site (see Table 1a and b). Thus a further analysis by a neutral party seems needed to establish which is correct.

Table 1. Summary of blood lead coefficients (with approx s.e.'s) found by Dupont and Schwartz in NHANES II and Pocock et. al. in British study.

	Systolic B.P.	Diastolic B.P.
a) Schwartz (NHANES II)		
males 21-74		
not site-adjusted	5.23 (0.95)	2.96 (0.62)
site-adjusted (all)	3.23 (1.0)	1.39 (0.66)
b) Dupont (NHANES II)		
males 12-74		
not site-adjusted	3.43 (0.9)	2.02 (0.6)
site-adjusted	1.95 (0.9)	0.36 (0.6)
c) Pocock et. al. (BRHS)		
all males		
not adjusted for town	0.675 (0.6)	-0.063 (-)
adjusted for town	2.089 (0.7)	1.809 (0.61)

1.2 The analyses of NHANES II data by Pirkle et. al. (1985 Am. Journal of Epidemiology) and the later analyses by Schwartz indicate that the blood lead effect holds up against an impressive battery of controls, excluding site: see Table 2, taken from Schwartz's (1985) paper to the Royal Society of Canada Commission on Lead in the Environment. Dupont's analysis also indicates a highly significant blood lead coefficient when site is not adjusted, although the size of the effect is smaller.

1.3. The lead coefficients on systolic blood pressure (SBP) and diastolic blood pressure (DBP) attenuate considerably when site is controlled, to levels that remain significant in the Schwartz analysis but are below significance in the Dupont analysis. (See Table 1). Thus the question of whether to adjust for site is important.

1.4. On the whole I support Dupont's argument that site should be adjusted. Schwartz argues against adjustment, noting that adjustment for site may decrease the power to estimate the effect of blood lead and (more seriously) increase the attenuation in the blood lead coefficient caused by measurement error. He also notes that specific confounders have not been suggested to account for the between-site blood lead effects.

However, if the relationship is causal we would expect to see an effect within sites. The standard error of the blood lead coefficient does not appear to change much when site is controlled, suggesting that little power is lost. Ware's comments correctly emphasize the dangers of inferring causal relationships from association between variables based on areal epidemiological data; the fact that specific between-site confounders are not known does not mean that they are not there, given the large unexplained variance and current state of knowledge. The attenuation effects from measurement error are a problem, but equally measurement error of measured potential confounders may bias the lead coefficient in the other direction.

1.5. I do not find Pirkle et. al.'s extrapolation of their results from NHANES II to determine coronary events prevented by blood lead convincing. The large number of debatable assumptions required to do this, leading to the inference that removal of lead will cause a decrease in blood pressure, which will in turn cause a reduction in coronary events, have been noted by other reviewers. Even if all these assumptions are granted, the choice of NHANES II blood lead coefficients, unadjusted for site, is highly questionable, both for the reasons noted above, and because they are markedly higher than those found in the large Pocock et. al. study. If any such material was retained in the EPA report, then I would argue strongly that all assumptions should be clearly discussed, consensus coefficients adopted, and measures of uncertainty included in the projections.

1.6 The restriction to White Males Aged 40-59 in the Pirkle et. al. analysis seems to me to have advantages, although the fact that both BP and lead are correlated with age for the whole sample can presumably be taken care of by age adjustment. The explanation of the restriction to ages 40-59 in the EPA report (page 4 lines 3-4 from bot) is misleading.

1.7 I am puzzled that the Schwartz analysis lacks a site-adjusted analysis of systolic BP for the 40-59 Age group. I hope this is not a case of selective reporting.

1.8 Unlike some reviewers, I do not find the results of the Pocock et. al. study necessarily in conflict with those of NHANES II. Table 1 a) and c) indicate marked differences in the blood lead coefficients between studies when site is not controlled, but reasonable consistency between the Schwartz and Pocock et. al. coefficients adjusted for site, particularly when sampling errors and the differences in control variables are taken into account. This consistency, and the divergence of the coefficients unadjusted for site, lends further support to the use of site-adjusted coefficients. The coefficients point to some effect, but of much lower magnitude than that found by Pirkle et. al.

1.9 More comparative analyses of NHANES I and NHANES II with respect to the effect of nutritional and dietary factors on blood pressure would help to determine the reliability of NHANES II for assessing the relationship of blood lead and blood pressure. In particular, the effects of dietary and serum Vitamin C seem to merit more attention. Data from the NHANES I study reported by McCarron et. al. (Science, 1984, Table 2) suggest that Vitamin C intake is negatively related to blood pressure. Data from NHANES II reported in Pirkle et. al. (Table 2) suggest highly significant positive relationships between dietary Vitamin C and systolic and diastolic blood pressure; a significant negative relationship is reported for serum Vitamin C for diastolic BP only. These analyses are not directly comparable, since covariates are not adjusted in the McCarron study. However, it would be useful to ensure that the two studies yield similar results when parallel analyses are run.

2. Comments on Joel Schwartz's reanalysis of Oct 2, 1985.

2.1 The main focus of the analysis is to develop refined controls of degree of urbanization, using the strata and city/smsa variables in NHANES II. Although interesting, the analysis does not resolve the discrepancies between the Schwartz and Dupont analyses, which were the basic concerns of Royall and Ware.

2.2 The analysis reads a bit like an attempt to scramble an egg without breaking it. Thus Schwartz states that "the use of location as a classification variable creates some difficulty since location is collinear with lead. This will...create substantial increase in the bias of B." However, location is not

"collinear" with lead, since there is variability of the lead levels within location - in fact the standard errors of the lead coefficients do not appear to change much when site is controlled, indicating little loss of power. Control for location will reduce the lead variability, but any potential confounder positively associated with lead with have this property. If the bias issue is used to rule out controls that are associated with lead, then the analysis is pointless, since association with lead is a necessary condition for a control to affect the lead coefficient.

2.3. Schwartz is right to raise the issue of attenuation of the lead coefficient due to measurement errors in lead blood determinations. The bias does become more important as the residual variation of lead levels is reduced by the introduction of potential confounders.

2.4. On the other hand, his analysis of measurement error is one-sided since he does not consider errors in the other independent variables. Measurement error in confounders will bias the blood lead coefficient the other way. For example, hypertension medication might be a confounder, since it has increased over the period of the study. I suspect this variable is included in Schwartz's analysis simply as a dummy for presence or absence. However dose and type of medication may be important. Thus this variable has "measurement error" which might make the lead coefficient look too high. A full analysis of this issue requires assessment of the impact of error in blood lead and potentially important confounders.

2.5. Schwartz argues that "there is a significant downward bias in the t-statistic that needs to be at least recognized". Ascribing bias to t-statistics is hazardous. The effect of measurement errors in the dependent variables reduces power, that is acts like a reduction in sample size. The effect will be to reduce the t-statistic if a real effect is present, but not if there is no effect. Thus attributing bias to the statistic begs the question of whether a true effect is there. For example, if the significance is marginal (say P=0.08), it is tempting but false to argue that t (and hence significance) would increase if the sample size was increased (or alternatively if measurement error was reduced). Schwartz might believe that his analyses have demonstrated a significant effect, but this is in conflict with Dupont's analysis controlling site. On the whole I would restrict the discussion of bias to the size of the coefficient itself.

2.6. A minor issue concerns Schwartz's argument that the sample design increased the collinearity of the PSU's with the outcome variable. Stratification on the outcome measures will not increase this collinearity unless extremes of the outcome variable are oversampled: was this done? Stratification does not distort the distribution of the sampled measures, but rather reduces the sampling error of estimates.

2.7. It would be useful when conducting analyses with a variety of alternative controls to present a summary table giving the size of the lead coefficient classified by each set of control variables used. This comment applies more to Schwartz's earlier analyses, where I would like confirmation that the minor percentage changes reported in the lead coefficient between analyses are not cumulating to form a relatively substantial effect. (I am not saying that I think this is happening). I would prefer quantitative summary tables to repeated statements that the coefficient didn't change much. Qualitative statements have a tendency to be subjective; for example, Schwartz states that control for location and stratification in one analysis has "very little impact" on the blood lead - blood pressure relationship. However, the lead coefficient for diastolic BP changed from 2.956 to 2.262, a 31% decrease which is more than a minor change.

TABLE 2 Variables Included in the Stepwise Regression Analyses

age *	dietary iron +
age-squared *	dietary vitamin A +
body mass index	dietary vitamin C +
dietary sodium +	dietary thiamine +
salt shaker sodium	dietary riboflavin +
dietary sodium X salt shaker sodium	dietary niacin +
	serum cholesterol+
dietary potassium +	serum vitamin C +
dietary sodium - potassium ratio	serum iron +
	serum transferrin saturation
dietary calcium +	serum zinc +
dietary phosphorus +	serum copper +
dietary protein +	serum albumin +
dietary fat +	hemoglobin +
dietary carbohydrate +	red blood cell count
dietary cholesterol +	ethanol consumption / week +
dietary saturated fatty acids +	cigarettes smoked / day
	total dietary grams +
dietary oleic acid +	total dietary calories +
dietary linoleic acid +	cigar or pipe smoking
blood lead+	

* forced into each regression to remove any possible age effects on blood pressure.

+ the natural log and squared transformation of these variables were also included in the stepwise regression.