E. I. DU PONT DE NEMOURS AND COMPANY
Haskell Laboratory for Toxicology
and Industrial Medicine
P.O. Box 50, Elkton Road
Newark, Delaware 19714-0050

DU PONT CENTRAL RESEARCH AND DEVELOPMENT

cc J.E. Leemann
C.E. Reinhardt, M.D.
D. Wegman, M.D.

March 24, 1994

TECHNICAL FILE
C-1516

Karen I. Bolla, Ph.D.
Assistant Professor of Neurology and
Psychiatry and Behavioral Sciences
Department of Neurology, Room 122-B
Francis Scott Key Medical Center
4940 Eastern Avenue
Baltimore, MD 21224

Dear Dr. Bolla:

Du Pont's Epidemiology Review Board (ERB) has reviewed your draft manuscript entitled "Clinical Course of Neurobehavioral Functioning after Chronic Exposure to Organic and Inorganic Lead." Enclosed is a composite review assembled from Board members' comments by Dr. David Wegman.

We hope you find this critique helpful. The ERB members ask that you send a copy of your final submission as well as the final manuscript, assuming the manuscript is accepted.

Thank you for your time and consideration.

Sincerely,

Judy Walrath, Ph.D.
Senior Epidemiologist

Attach.

N 29304

WORK ENVIRONMENT
Department of Work Environment
University of Massachusetts Lowell

March 21, 1994

Judy Walrath, Ph.D.
Epidemiology Section, N-11510
Haskell Laboratry CR & D
P.O. Box 50
Newark, DE 19714

Dear Judy,

Attached is a composite critique of the draft manuscript: "Clinical course of neurobehavioral functioning after chronic exposure to organic and inorganic lead" by K Bolla, L Kahler and J Rignani. I hope you find the comments useful. I think the Epidemiology Review Board members would be interested to see the final submission as well as the final accepted manuscript (assuming it is accepted). In this way we could see how they chose to respond to our critique as well as to see what changes resulted from the journal's peer review process.

Sincerely,

David H. Wegman, M.D.
Professor and Head

N 29304.01

One University Avenue Lowell, MA 01854 USA Telephone: 508-934-3250 • Fax: 508-452-5711

DuPont Corporation Epidemiology Review Board

Review of

"Clinical course of neurobehavioral functioning
after chronic exposure to organic and inorganic lead"

by

Bolla KI, Kahler L and Rignani J

This manuscript reports results from the follow-up of a small group of workers (23) from among 222 who participated in an earlier cross-sectional study of workers exposed to mixed inorganic and organic lead. The follow-up study group was comprised of 17 who volunteered for re-testing and 6 who were "advised to do so" by the original study team. Not all 23 workers completed the full battery of tests for the follow-up exam. e.g. only 13 workers completed the Q16 (Scandinavian Questionnaire Sixteen) at follow-up.

The majority of workers showed no significant change in performance (as defined by the authors) over time on objective neurobehavioral tests, although there was an increase in the reporting of symptoms. Significant improvements or deterioration is reported for roughly the same number of tests; three versus four respectively. All four of the tests that showed a decline over time were tests involving psychomotor speed or a rapid motor response. Of note is that 9 of the 17 workers with normal neurobehavioral evaluations in March 1990 had abnormal evaluations at follow-up, a finding which the authors attribute to a psychosocial mechanism rather than a toxicologic mechanism although there is no direct support provided for this interpretation of the findings.

Study Design: As the authors emphasize, the population is small and self-selected, making generalizations difficult. Only six of the subjects in the report were advised to be re-tested. Without knowing the total who were so advised, the reader cannot determine whether these workers are in any way representative of the entire group for which re-testing was recommended.

A similar problem concerns the 17 workers who were not specifically counseled to be re-tested but nonetheless chose to be (perhaps the appropriate questions were not asked). It is quite likely that the workers' perceptions of deterioration of neurobehavioral functioning were the motivating factors. One would therefore expect to see a worse performance on re-testing of these workers.

The authors should indicate how many of the 222 total workers tested in the baseline survey were offered the re-test opportunity. It is assumed that the two groups

of workers studied are incidental samples of two groups that comprise the baseline survey. It would be of further assistance to the reader if the results from the baseline for all workers in each of the two groups (those recommended for follow-up exam distinct from the others) were also provided. Finally, no information is provided as to how the 23 participants compared with the remainder of the 222 workers with respect to lead exposure.

In brief, selection bias is a distinct possibility that has not been addressed adequately. Selection bias, in general, may be more of a concern in an epidemiologic concerned with assessing etiologic relations than in a clinical case series of natural history of disease. Nevertheless, the inferences drawn from this study could be erroneous.

Lead Exposure: As currently presented, it is not possible for a reader to associate exposures to lead of individual test subjects with their test results. The authors report their interpretation that lead exposure levels were not associated with differences in test scores for the two test sessions, but the reader has little or no sense of the levels of exposures involved with test score differences.

The methods section of the manuscript contains a relatively brief description of exposures at the bottom of p4 and top of p5. While this description includes reference to exposure zones, exposure zones are not used to address exposures of subjects in either the narrative or the tables of the current manuscript. The exposure and dose measures which are addressed in the current manuscript (Table 2) are exposure duration, lifetime weighted average blood lead level, and lifetime peak blood lead level; none of these indices require reference to the 29 exposure zones.

While information on duration of exposure is presented in reference 6 (Am J Epid [1993] 137:1006-1021), blood lead levels presented in the reference are only a single value for the lifetime average for all 222 study participants and a single lifetime average value for all 98 nonparticipants. Reference 6 does contain the following statement: "Blood lead and zinc protoporphyrin data were largely available for only the last 3-5 years, and thus were not adequate to estimate the cumulative dose in workers employed for longer durations. Data also suggest that blood lead is an inadequate biologic measure for organic lead exposure (24)."

In the current manuscript there is very brief discussion of the pertinence of lifetime weighted average blood lead levels and lifetime peak blood lead levels to the issue at hand. There is very little context, however, in the current manuscript or in reference 6 for the interpretation of the blood lead data which are presented (Table 2).

The "...important temporal trend in exposures..." identified in this manuscript's discussion of study subjects' exposures to lead is not further discussed in this manuscript nor is it examined in reference 6. Some discussion of its importance and,

because the authors consider it to be important, some presentation of data illustrating the trend and the magnitude of change, would be helpful to readers.

The report that workplace exposures decreased in the period September 1989 to May 1991 is not documented elsewhere in this manuscript. Documentation was not found in reference 6. The only pertinent comment found in reference 6 is: "Workers reported that personal protective equipment was not frequently used before August 1989, and therefore, such equipment was not considered in deriving individual exposure estimates." If a decrease in exposures is important in the interpretation of findings, some documentation that exposures actually decreased for the study subjects may be in order.

Data Analysis: The approach to data analysis has some important deficiencies. First, the categorization of "better," "same," and worse" is based on a confusing and questionable algorithm. As stated on p6, worse performance on re-testing means that the change in test results from (t2 - t1) was in a negative direction and more than one standard deviation below the mean score for the original study group of 222 workers.

This could be interpreted in two ways. Suppose the mean score for the full cohort was 100 and the SD was 10. Then one SD below the mean would be a score of 90. One way to interpret the categorization is that if a worker's score fell below 90 it is "worse". This would require anything from a trivial (e.g., 90 to 89) to a dramatic and probably unrealistically large change (e.g., 120 to 89. Furthermore, it would be impossible for a worker, whose initial score was less than 90, to be categorized in the "worse" group. The other interpretation is that if the change in score is > 10 - no matter what the absolute score - then the change would be categorized as "worse". The latter definition is insensitive to the initial score (a decline from 180 to 170 may have a very different meaning than a decline from 80 to 70), and may merely reflect regression to the mean. For example, the apparent decrements in manual dexterity (Table 3) are consistent and may be important, but this cannot be determined because of the method of analysis. There are much better ways to analyze repeated measures data, and some of these maintain simplicity of presentation without jeopardizing validity.

In either event, a clearer statement of the method for categorization needs to be presented. If the latter categorization is the correction one (as we believe it is likely to be), then the statement of the criterion ought to indicate that a worker was placed into the "worse" category if the change in score (t2 - t1) was greater than or equal to one SD of the entire group of 222 lead workers (e.g., if $\Delta > 10$).

In addition, given that the distribution of the test results in the initial survey of 222 workers is critical to the definition of the outcomes, the authors ought to present

a table with the means and SD of each test in the baseline survey so that the reader has an idea of how big the SDs are.

Finally, the authors need to consider the problem that is present due to the range in length of follow-up of the 23 subjects. The reported follow-up varied from 1 to over 2 years, but no mention is made about what is known concerning the expected change in each test due to aging in normals. More importantly, it should be made clear whether or to what extent the greater changes that are reported occurred among those with the longest time between tests.

Results and Interpretation of Findings: In presenting results (p8) the authors call attention to the increased symptom reports on Q16. They note, but do not specifically address the findings on the last item which is included "to measure over-reporting of symptoms." As noted earlier participation may have been determined by the workers' perceptions of deterioration of neurobehavioral functioning and therefore the authors might have expected to see a worse performance on re-testing of these workers. When comparing the prevalence of symptoms from the Q16 questionnaire (Table 4), it should be made clear that the percents refer only to the 13 who participated in the symptom portion of session 2.

Throughout the presentation of results, attention should be clearly directed to differentiating the two components of the follow-up group. The data should be presented in each table so that the reader can distinguish those who were recommended for re-test (6) from those who self-referred (17).

In interpreting the findings from the follow-up results on the 23 subjects, the authors conclude that "for the majority of workers, little change in neurobehavioral functioning occurred over time". However it is not obvious to the reader that a change in SD of 0.9 (below their 1 SD criterion) in the period of a single year is trivial even though it would be classified by these authors as the "same".

For pulmonary function, for example, a recent cross sectional study found the mean FEV_1 was 3.64 liters with a SD of 0.74 (approximately 20% of the mean). Since a healthy individual can be expected to lose FEV_1 at the rate of approximately 1% per year (0.037 liters per year in the example), this would mean that in a two year period a subject would be expected to lose only one tenth of an SD (0.074 liters). If a subject were to lose more than *one* SD, that would mean that he/she had lost more than 20% of lung capacity in a single year. This would be cause for alarm.

This example, based on an outcome with well characterized behavior in populations, illustrates how much relevant information has not been provided about the neurobehavioral test outcomes. For FEV_1 a reasonable criterion for "worse" annual change might well be half or even one quarter of the SD. Without some similar information about for the measures used in this manuscript, it is not at all obvious that

a change of <1 SD is a reasonable definition of abnormal for these neurobehavioral tests. The authors need to provide information relevant to this definition of abnormal change. The absence of such information further complicates interpretation of the findings.

The main conclusion reached from this analysis is that changes indicative of decline in neurobehavioral performance over time are more likely to be caused by psycho-social factors than by lead exposure. This may be true, but the study was not designed adequately to evaluate the effect of lead. The implication that the workers may have assumed a "sick role" that resulted in poor test performance is speculative and could be interpreted in a variety of misleading ways: reducing lead exposure has effectively curtailed neurobehavioral declines, the workers are malingerers, etc.

It is surprising that the authors chose not to show any of the findings related to lead exposure, age, or other factors. The reasoning (pδ) is that there were no observed statistically significant odds ratios. This is a poor decision in view of the small study size and low statistical power. The associations with lead exposure are the most important findings; not showing them is unfortunate.

While this study is severely limited by the small and nonrepresentative sample of workers available for follow-up, and the incomplete data available at follow-up, one interpretation of the findings could be that anxiety about health and well being may have resulted in the findings. This possible interpretation deserves further study in larger and more representative cohorts and raises questions about the effects anxiety may have had on the results of the testing done at the time of the initial cross-sectional study.

Recommendation: This paper needs substantial improvement, in the areas of data analysis, presentation of results, and interpretation. All of these aspects could be remedied, but problems of selection bias cannot.

Page 2, Abstract, line 12. "No measure of mixed organic and inorganic lead exposure (e.g. duration, intensity),..." The manuscript does not describe or discuss exposure intensity.

Page 4, line 1. This appears to be a statement rather than a question.