3001 Derby Street, Apartment C
Berkeley, California 94;705
July 30, 1995

Ken Mundt, Ph.D.
Applied Epidemiology, Inc.
P.O. Box 2424
Amherst, MA 01004

Dear Ken,

Through the vagaries of the Postal Service, I received you package on July 18, over a week after you mailed it Priority Mail. Unfortunately, this has been typical of other packages I have received Priority Mail from Massachusetts--they have all taken a week or more! Then, I had business out of state. So, this weekend was the first I could get to your proposal. My comments are on the following pages. I will both mail and fax these. I hope these are useful. If you have any questions, please call me at (510) 643-0289, or fax me at (510) 642-5815. Alternatively, my home phone/fax number is (510) 843-4329. Good luck with the study!

Sincerely,

S. Katharine Hammond, Ph.D.

Comments on the Protocol for CMA Vinyl Chloride Mortality Study
Submitted to Kenneth A. Mundt, Ph.D.
from S. Katharine Hammond, Ph.D.
July 30, 1995

Overall, this is a very good protocol for a complex study. These comments are focused on the exposure assessment issues, although a few other areas are mentioned as well. All of these comments are intended only as suggestions. However, even if some are beyond the scope of this study, you may be able to determine the limits of the data, or what further analyses might be possible in future studies. Although "exposure assessment will not be enhanced as part of this investigation," you may be able to do more with the data in hand than you had planned, but what is "in hand" is not clear.

Exposure Assessment Issues

Your review of the literature might comment on the quality of the exposure assessment in the studies you discuss, and how this might limit interpretations of the findings of each study.

Although on p. 15 you list data tapes with "date/coded exposure to vinyl chloride," the inventory of boxes received from ENSR, from Addendum I, does not list these per se; are they part of the individual subject files? Is there a job exposure matrix for each plant which you also received? Or, do you have some condensed exposure code? The uncertainty arises again on p. 27, first paragraph, in the discussion of "accumulating person-time by various risk factors of interest, *provided a job exposure matrix and standardized work history are available*." [emphasis added] Are they available for this cohort? Many of the following comments assume that they are available.

The protocol is confusing as to what you will be using for exposure variable(s). On p.17, first paragraph, you list two definitions: 1) exposure for each job and each job location was defined on an ordinal scale as "low, moderate, high;" and, 2) each job is classified as exposed or not exposed; the paragraph concludes with the statement that "the current update will review, and if appropriate, employ the second definition of exposure, given the classification problems associated with the original method." What were these problems? Were they worse than losing the distinctions between those with only trivial exposure and those with distinct, higher, exposures? On p. 23, 3rd paragraph, you indicate that you will use the 3 point ordinal measure of degree of exposure, in addition to the dichotomous measure and duration of exposure. What data are available to evaluate exposure to VCM only, PVC only, and to them both? Will all three predictors of exposure (dichotomous, three levels of exposure, and duration of exposure) be used for each of these qualitative aspects of exposure? Will intensity of exposure be combined with duration, comparable to a cumulative exposure index such as ppm-years?

Assuming you do have job histories for each worker and a job exposure matrix for all jobs at all plants in all years of operations, the question arises just how to use this exposure data in the epidemiological analysis. Should all those with any time in "high" exposure jobs be placed in the "high" exposure group? How should one incorporate intensity of exposure with duration of exposure? A few suggestions follow:

1) Investigate the distribution of exposure, that is, determine the number of people ever in each of the exposure levels and the number of person years accumulated in each of the three levels of exposure. You may decide to combine the top two levels of exposure for some analyses.

2) Examine the industrial hygiene (if any) or other data to determine the relative intensities of exposure in the three levels. Do not assume that they are 1, 2, 3, i.e., that the highest level is three times higher than the lowest level. In fact, most data indicates that the ratios are much greater. In the Semiconductor Health Study, we divided tasks into high and low intensity, with a ratio of 50, based on measurements made of the tasks that were classified as high and low; you might also refer to some of Jerry Lynch's work, in which he also reports large factors between ordinal exposure levels. This becomes important if you plan to accumulate exposure in the different intensity categories over an individual's work career.

3) There is a risk that combining all exposed workers into one group will dilute the effect of the agents of interest by combining those with trivial exposure with those few with sufficient exposure to have a health impact, e.g., all steel workers vs. coke oven workers. Consider analyzing the dichotomous data a few different ways: in addition to exposed vs. not exposed, try exposed above some threshold vs. unexposed or exposed below the threshold. The choice of thresholds will depend on both hypothesized mechanisms and the distribution of your data (see point 1 above)

Other Issues

How will you account for the healthy worker effect if you use all US males (or all regional males) as your comparison group?

p. 11,1st paragraph, 3rd line: You might consider adding the following qualifier, as italicized: Health hazards *known to be* associated with this industry

p. 17. Explain which are the "Group A," "Group B," and "Group C" states, and what is the meaning of this categorization.

p. 21 You might consider using the data from the 1950 census for 1950-1955, the data for the 1960 census for 1956-1965, and the 1968 data for 1966-1968.

p. 30 2nd paragraph: How will you report the results for outcomes with fewer than 2 expected deaths for which there are more observed events?

p. 34 Do you have a plan to notify the study subjects, that is, the surviving workers and the families of the deceased workers? Note page 15 of the Guidelines for Good Epidemiology Practices.