UNIVERSITY OF WASHINGTON

School of Public Health and Community Medicine
Department of Environmental Health, Box 357234
SEATTLE, WASHINGTON 98195-7234

Tel.(206)543-2052
Fax(206)685-3990

July 14, 1995

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

RE: "Protocol for CMA Vinyl Chloride Mortality Study"

Dear Dr. Mundt:

I have reviewed the protocol and have some comments and suggestions for finalizing the protocol.

General Comments

The study objectives and procedures are very clearly presented, following accepted epidemiologic methods. The proposed update of the Wong study should, by virtue of a longer follow-up and additional deaths, provide much needed information about potential hazards related to VCM and other chemicals in this industry. It seems evident that your group has given an appropriate level of attention to issues of data quality and completeness, which will certainly be necessary for producing a valid study.

The approach to the data analysis, basically SMR comparisons against external rates as an initial screening for effect estimates, followed by in-depth analyses relying on internal risk comparisons, is sensible. This is standard practice in such studies, and usually provides an adequate characterization of mortality risks.

The main areas that could use further clarification in the protocol concern the amount and detail of exposure information, and how these data will be incorporated into the analysis. In particular, the description of the original Wong study indicates that a fairly crude exposure classification system (exposed vs. not exposed) was adopted. As correctly mentioned, this scheme is only minimally applicable for estimating dose-response relations (more specific comments on this issue will follow). I think it would be of considerable value to append to the proposal a copy of the most recent Wong report to permit a comparisons of existing knowledge with what should be obtained from the update. The proposal would also benefit from some description (brief) of the plants and the study populations to orient uninitiated readers. There is a listing of boxes of study materials provided by ENSR, but the meaning of this obscure to me.

There is some passing reference to the desirability of performing nested case-control studies for diseases of prior concern (e.g., liver cancer) and other diseases for which consistent excesses are seen. More should be said about how this would proceed, i.e., how diseases for case-control studies would be selected, control sampling strategies, collection of additional data (e.g., industrial hygiene data, smoking), if this is envisioned.

On balance, this is a well formulated proposal that outlines the plans for what should be a scientifically valid investigation.

Specific Comments and Suggestions

1. A considerable effort will be given to the construction of external reference rates, apart from those currently available by standard lifetable programs (e.g., Marsh). Is this effort really necessary? The University of Pittsburgh can provide state- and county-specific mortality rates that are useful; that has been my experience.

2. Unknown race my pose a problem for some diseases. Is this the case for brain cancer or liver cancer?

3. As mentioned earlier, the quality of the exposure data will be a critical determinant of the study's ultimate value. The dichotomous classification and subjective rating schemes used previously do not appear to be rigorous enough for meaningful dose-response estimation. It might be worth including a feasibility study to determine the availability and quality of historical exposure data that could be exploited, perhaps in subsequent nested case-control studies. It is likely that exposure information is generally sparse, but there may be some plants among the 37 included in the study that have reasonably good data. If this were true, then subcohort analyses of these plants might be explored.

4. The diagnoses of liver and brain cancer from death certificates may be uncertain (e.g., confusion with metastatic cancers of other primary sites). The possibility of diagnostic errors will need to be addressed.

5. Some planning for nested case-control or case-cohort studies would improve the study, even if the planning were limited to a determination of feasibility.

6. The discussion of statistical power (p20) is not especially enlightening. The study needs to be done; thus, whatever power might be attained is really a moot point.

7. If intense VCM exposures are neurotoxic, then there may be a healthy worker survivor effect that could attenuate observed associations. This might be mentioned, although analytic methods to rectify the problem are very complex and may not yield interpretable results in the absence of quantitative exposure data.

8. Other analytic features that could be considered are induction/latency analyses (exposure lagging and time windows), and stratification of person-time by actively employed status.

I hope that these comments are clear and useful. Please do not hesitate to contact me if you need clarification of these remarks or additional comments.

Sincerely yours,

Harvey Checkoway

Harvey Checkoway, Ph.D.
Professor of Environmental Health
and Epidemiology

SL 108234

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.

AMENDMENT I.

RESPONSE TO REVIEWS BY DRS. CHECKOWAY AND HAMMOND 8/1/95

The comments of reviewers Dr. Harvey Checkoway and Dr. Kathy Hammond may be found in *O. Protocol Review and Approval Sign-off Sheet*. The specific comments and suggestions of each reviewer were evaluated and incorporated into the body of the preceding text. However, one recurring theme was evident in each review, independently - the need for a better assessment of exposure. While such an assessment is beyond the scope of contract requirements, we do agree that an exploration of existing sources of data is warranted, and would aid in evaluating the feasibility of further exposure assessment. Our ability to perform such an exploration and subsequent assessment is greatly hampered by the fact that much of the data is currently lost or unaccounted for. The eventual securing of this data will allow us to conduct a feasibility study to determine the availability and quality of exposure data, perhaps leading, eventually, to a nested case-control study.