By: Carl E. Hopkins, University of Oregon Medical School

Statisticians and public health workers alike are in debt to Dr. Dorn and the National Cancer Institute for this painstaking contribution to our knowledge of mortality and its relation to smoking. It is truly remarkable that so little is still known about causes and effects of such a widespread human activity as tobacco use. Physicians, philosophers and poets have for centuries speculated, eulogized and moralized on the subject, without adducing substantial evidence, until Lombard, and separately, Pearl, made in the late 1930's some epidemiologic studies of the effect of smoking on mortality and longevity. The results were not shocking, and it was not until the 1950's when a sharp rise in lung cancer deaths became noticeable here and in England that intensive exploration of smoking as a possible health hazard began. In retrospect it is easy to see that just as in much exploratory experimentation in the laboratory, a main effect had been completely obscured by excessive dilution until a sufficient concentration was reached to make the main effect stand out above the noise level.

Now that we have a main effect standing out clearly, it is also remarkable, tho not unprecedented, that rather unusual standards of credibility are being required for the conclusions of such epidemiologic investigations. While arising out of diverse motivations and confused understandings of the statistician's role, these unusual requirements of "proof" have been a stimulating challenge to the professional statistician and produced some worthwhile introspection on his role and his method.

If we view his role as that of making some sort of rational decision on whether to adopt or reject some incompletely specified but farreaching public health program, then it is quite clear that he does not yet have values for all of the parameters of the pertinent decision functions. For that matter it is not clear that he even has the necessary tools. But if we view his role more narrowly as that of evaluating the credibility of inference drawn from observations, then it would appear that his usual tools of probability and sampling theory as applied to statistical epidemiology may be presumed to suffice here as they have in so many other similar problems where direct experimental test seemed impossible. I prefer to speak from this latter viewpoint, and will therefore discuss Dr. Dorn's contribution only as to the inference he draws from his observations, viewing wider implications, however important they may be to the public health, as irrelevant in a statistical discussion.

The National Cancer Institute study was ingeniously conceived to eliminate certain ambiguities remaining in earlier studies due mainly to their retrospective design, their susceptibility to bias in selection of smokers and non-smokers, and their size limitations. While being motivated by interest in the lung cancer problem, the study was wisely broadened to cover mortality of all kinds in relation to smoking. Veterans provided a large, defined and apparently suitable population for following in a prospective study. The study as reported shows every evidence of care and great technical skill in planning, execution and interpretation. Few statisticians could really wish to undertake such a task, but most could certainly feel proud of the job reported here. For the first time in all the welter of shaky claims and counter-claims the relations of smoking to overall risk of dying and to risk of dying of particular disease stand out clearly as facts against which hypotheses may be tested.

It should be added quickly that the study says nothing about causes or mechanisms for explaining these facts. This must be done eventually by the epidemiologist and the pathologist who must integrate these and all other known facts into a plausible and preferably testable explanatory hypothesis. Several of these are on the market now, and facts from this study should help in selecting from among them.

Yet there are limitations to these facts. Some of the old bugaboos that hobbled earlier studies are still with us.

 Retrospective aspects. Altho prospective since 1954, the crucial separation into experimental and control groups by smoking history is in some senses retrospective. The subjects of the study are U.S. uncolored males, who (a) became veterans of the armed forces before 1940, (b) continued their life insurance in force until 1954, (c) survived until 1954, and (d) responded to a smoking history questionnaire in 1954.

The first 3 of these selections are obviously influenced by such host factors as health status, including mental health and personality, social and occupational class, etc.; all of which could be associated with smoking and non-smoking. Since we end up with a reference population only about 16% of which report as non-smokers, a little arithmetic will show that rather moderate biases in the selection of smokers and nonsmokers into these successive stages of the prestudy history could result in a disproportionate overloading of extra-healthy persons into the non-smoking part of the 1954 study population, thus giving non-smokers a superior survivorship due not to non-smoking but to health and social factors that made them non-smokers. Until we know more about the social, psychological and physical factors in the individual person that select him into the non-smoking, smoking and heavy-smoking classes, we cannot be certain that these mortality data do not at least partly reflect the mortality prognosis of the groups prior to their selection into the smoking classes. These selective effects may be presumed to have been washed out by the regression phenomenon in the interim from beginning of smoking until entry into this mortality study, but we have no way to know how much to allow for this.

2. The memory problem. Here, as in Kinsey's studies, the question is not whether the questionnaire obtained reliable, repeatable responses, but whether the responses were <u>valid</u> in the sense of corresponding with objective fact. Some skepticism is required on the accuracy of the memory instrument in a subject as charged with deep psychological content as smoking habits.

- 3. What of the non-response bias? It is presumed that the final report will explain what effect the 16% non-respondent group could have had on the outcome.
- 4. What of experimental artifacts? Did the policy holders know they were queried thru their life insurance company? Could this have influenced their responses? Could they have feared to lose their policies or to have premium rates changed as a result of their replies?
- 5. To what larger population can these results be generalized? They were obtained from U.S. uncolored males, somewhat upper class, with 40% better than average mortality experience. What can the epidemiologist do with this?

These are not detractions. But they are questions and limitations that must be kept in mind in weighing and using the evidence of this study. If the study were to be viewed in isolation one would have to be suspicious of dangerous biases possibly lurking in the background. Its credibility tonite rests not only on the evident craftsmanship of the work itself, but also and perhaps mainly on agreement at crucial points with numerous other independent studies subject to <u>different sets</u> of biases.

In sum, we are indebted to Dr. Dorn and his coworkers for their significant contribution in confirming previous but less well established facts, and bringing out new detail on kind, amount and length of smoking history, as related to specific cause of death. We are another long step ahead with facts which will eventually resolve controversy.