

Role of Human Population Study Centers in Studies of Cancer Etiology

WILLIAM HAENSZEL and ROBERT W. MILLER, M.D., Dr.P.H.

RECENTLY interest has quickened in the use of human population study centers as sources of data for medical research. The idea is not new. An early example of population studies was the series of investigations in Hagerstown, Md., initiated by the Public Health Service in 1921 (1). Other such studies have been supported in part by National Institutes of Health research grants. The Framingham, Mass., survey of heart disease (2) and the Veterans Administration-National Cancer Institute study of smoking and health (3) were National Institutes of Health intramural activities.

To assess their role in the support of human population study centers the NIH Division of Research Grants Study Sections concerned with grant applications sponsored the Conference on Health Studies of Human Populations at the University of Pittsburgh School of Public Health in November 1960 (4). The question now is what steps the National Cancer Institute should take in the further development of human population study centers.

Any review and appraisal of human population study centers is conditioned not only by events past but by the expectation of things to

come and so must be regarded as speculative judgments not unlike those involved in predicting the future course of business and market activities.

Humane and ethical considerations deprive investigators of the distribution of disease in man of many advantages conferred by experimental work with pedigreed animals, such as precise control of environmental and related factors, detailed and sophisticated comparative measurements in experimental and control animals, and the relatively large numbers of subjects available for experiments. Studies of disease in man must rely heavily on observations and on the substitution of statistical for experimental control. Situations, then, must be sought which improve the power of the observational method, usually through resort to more detailed cross-classifications of data and more elaborate determinations of the sequence of events, in order to extend the range of inferences permitted by the observational associations.

Many investigators look upon human population study center studies as a specialized activity which has its own techniques and discipline. In a statistical context, this distinction arises from recognition that statistical methods suitable for analysis of experimental data are often not appropriate for handling observational data.

The "human population study center" label has been applied to a number of different concepts so that it is difficult to obtain general agreement on all features of these centers. At the outset it would be well to distinguish between "populations for continuing study" and community-based "population center facilities."

Mr. Haenszel is chief of the Biometry Branch and Dr. Miller is chief of the Epidemiology Branch, National Cancer Institute, National Institutes of Health, Public Health Service. This paper was prepared for presentation at a meeting of the Consultant Panel on Biometry and Epidemiology of the National Cancer Institute on November 5-6, 1961. While the present text benefits from the comments of panel members and other reviewers, the authors assume responsibility for the views expressed.

The concept of populations for continuing study leads directly to considerations of longitudinal studies. The alternate concept of community-based population center facilities would cover a number of other study approaches, among them the use of center facilities to obtain general population control data for retrospective and cross-sectional studies.

Longitudinal studies are usually expensive and one should scrutinize any proposal calling for collection of data over a long-time interval to determine why a more conventional study as of a fixed cross section of time would not do. The simplest answer, which has justified the design of certain studies of rare events, is that a prolonged interval is required to accumulate sufficient numbers of observations to permit a meaningful interpretation. The increased number of events observed in longitudinal studies will often allow greater utilization of multiple classification of variables. Such utilization favors the study of the interactions of several variables on the biological responses in question and facilitates choice among possible models postulated to account for the underlying mechanism. "Interaction" is a general term and would cover such other obvious topics as familial aggregation and clustering by occupation, social class, and so on.

A more fundamental justification for longitudinal studies is that seriatim, long-term observations are needed to quantify "growth" characteristics of individual members of the study population. In more general terms the problem is one of estimating probabilities of the transition of cohort members from one defined state to another over a fixed time interval. The longitudinal approach also becomes essential if the trait measurements are thought to be unique to an individual and their relative direction of change with time the item of primary importance, with the absolute measures relegated to a secondary role (5).

Also, information as of a cross section of time is limited to well-recognized end points of morbidity or mortality diagnoses. Even when information on transitional states is collected as of a fixed time, it must be packaged in the form of prevalence rates. Under rather severe assumptions some deductions on transitional probabilities can be made from these rates, but

this leaves many uncertainties. Nothing can replace seriatim observations to establish precise estimates for transitional probabilities. Knowledge of the predictive value of antecedent events is essential to solutions for the family of "competitive risk" problems. This type of information, of which very little exists, would be useful to sort out alternate interpretations of observational associations.

To summarize by analogy, longitudinal studies can offer further insights into the submerged part of the iceberg: the transitional states and their associated probabilities which underlie the conventional observational end points.

Choice of Populations

In addition to the study objectives, some considerations which have governed the choice of human population study centers in the past are (a) unusual frequencies or distributions of specific diseases or exposures in the study population; (b) the presence of special diagnostic or medical care facilities (populations under comprehensive surveillance); (c) special systems which make followup relatively easy as, for example, insurance records; (d) certain demographic characteristics of the population (other things being equal, it is better to choose a community attracting in-migrants rather than one losing population through out-migration); and (e) the locale of interested investigators and institutions.

There has been a spirited division of opinion as to whether human population study centers should be conceived primarily as omnibus study resources within which large numbers of investigations, related or unrelated, can be developed, or as ad hoc studies with well-defined objectives. The omnibus concept has led to comprehensive community study programs in Tecumseh County, Mich., and in Alameda County, Calif. The Tecumseh County operations have been transformed during the initial 2 or 3 years into a study with more sharply defined objectives. The ad hoc concept is exemplified by the study of radiation effects among Japanese survivors of the atomic bombs (6,7), and by Morris' observations on physicians (8) and on London bus drivers and conductors (9).

Most investigators would not stress the "representative nature" of any human population study center, although some would place a very high premium (mistakenly, in our opinion) on this property. Any determination of what constitutes "representativeness" would be conditioned by the study hypothesis. Another reason for questioning the importance of "representativeness" is that interest in these studies would normally be directed to relative patterns and configurations rather than to absolute estimates of parameters descriptive of the general population; there is some reason to believe that generalizations about relative relationships hold true under a wider range of circumstances than those concerned with absolute magnitudes. Finally, the characteristics of any one source material become less crucial when inferences depend on collation of evidence from several sources.

Multidisciplinary Approach

For the professional staffing of the human population study center, the trinity of a physician-epidemiologist, a statistician-demographer, and a social psychologist will usually be the irreducible minimum. Other specialties will be required, depending on the nature of the inquiry. It is essential that there be responsible overall direction of the center, skilled in the exercise of a strategic viewpoint which can balance the commitment of money and personnel against anticipated study returns, in the context of numerous conflicting and apparently irreconcilable demands. The need for such direction cannot be overemphasized, for human population study centers offer many opportunities for diversionary efforts, the pursuit of which can lead to the defeat of any study plan.

Professional or scientific achievement alone does not confer supervisory skill; there is no sure way to identify persons with aptitude for strategic and tactical planning, unless they have proved themselves by past performance. No professional specialty has a monopoly on persons with these aptitudes, which are as difficult to identify as it would be to write a protocol on the conduct of a game of chess. Nevertheless, case histories of large-scale population-based inquiries which have run into difficulties because

of the lack of strategic and tactical planning can be advanced to document this need. Unless the project director has these administrative talents in sufficient degree, in combination with some other professional skills, no otherwise well-conceived set of studies is likely to prove productive and fruitful over the long run.

Source Materials

The data collected in the course of special investigations in population study centers should be collated with other information available from community sources. So great is the effort spent in developing rapport and cooperation for this purpose that there is much to be gained by using any human population study center as a vehicle for a variety of studies.

A set of difficult problems requiring close study and further research concerns the extent and scope of collecting routine intelligence and its organization into a record system and determination of how this information relates to present and future study objectives. In addition, the mechanics of collating routinely collected information for a defined population base need to be improved. The present clerical procedures for relating records to given individuals are time consuming and should be mechanized. The use of computers offers considerable promise, and methodological experiments in this area should be supported.

Analysis of Data

When seriatim observations are an integral feature of a study, the obvious analytical approaches to the data are variations of the actuarial method. These approaches can provide great flexibility in the manipulation of data in the form of single- or multiple-decrement tables. The difficulty with the prospective approach is that the relevant discriminating variables must have been included in the study design. If the investigator were so unfortunate as not to have made the right selection of study variables, no technique of analysis would be helpful.

Enough experience has now been gained with retrospective studies of diagnosed cancer cases and controls to convince most people that with

suitable precautions these studies will yield useful measures of relative risk. What may not be so widely appreciated are the technical advantages of supplementing the interpretation of prospective results by grafting onto them a retrospective approach to the same populations. A dual prospective and retrospective approach to the same population will go far to reduce the major risk of prospective studies—failure to identify the major discriminating variables in advance.

Study requirements will differ depending on whether the transitional probabilities to be determined are related to a stable, fixed trait, such as blood type, or to characteristics which change with time, such as vaginal cytology. For the fixed trait, a dual sampling approach may suffice. After assembling data for a specified numerator of cases, appropriate denominator values can usually be obtained by drawing appropriate samples of the population at risk. When the associations sought concern a dynamic trait, data for the denominators must be collected in advance on a cohort of individuals to be followed prospectively. In this situation the data collection requirements can involve not only a different order of magnitude but also a schedule for processing the collected materials which is quite different from the schedule called for by study of an association with a fixed trait.

General Comments

Studies of well-defined populations in depth clearly permit certain advantages over those available in other observational settings. Many future advances in cancer epidemiology will come from intensive use of these facilities. Nevertheless, the success of any human population study center undertaking is not automatically assured. Many problems remain to be resolved and a certain proportion of failures must be expected. A program of further support for human population study centers must be concerned with identification of promising study situations and the establishment of reasonable safeguards in making commitments.

Most persons could subscribe to the following criteria for judging human population study center proposals:

1. The proposal must incorporate at least one

study idea with merit in its own right; this would not preclude utilization of the human population study center for other studies at the same time or at a later date. Adherence to this requirement would sharpen operations and even improve the possibilities for later broad spectrum use of a human population study center. At worst, if events develop unfavorably, a specific objective will force a decision on a termination date to cut losses.

2. The desirability of using a human population study center would be heightened by the presence of unique features related to the study objective, which either complement or fill gaps in information available from other human population study centers. For example, if existing studies of this sort were concentrated in metropolitan cities of the northeast, a human population study center to investigate a farm population in a midwestern State might be encouraged. A certain amount of duplication in coverage should be encouraged since it may often be advantageous to replicate certain results in different human population study centers.

3. The assets in the form of available diagnostic facilities and provisions for long-term followup should be attractive. However, the mere presence of good supporting facilities is not sufficient reason to start a human population study center.

4. The character of the project direction should be carefully assessed, not only the specialty staffing pattern but, even more important, the overall scientific and executive qualities of the working, not the pro forma, project director. Judgment should not normally be made in the absence of identification of the proposed responsible director, who will devote the major part of his time to conduct of the project.

5. Since determination of transitional probabilities should be a central feature of many of these endeavors, any proposal should incorporate some provisions for diagnostic or screening tests with potential applications for the delineation of precursive states. In view of the investment required, support should be reserved for those studies which promise to break new ground in an epidemiologic-experimental approach to carcinogenesis.

Lack of trained personnel will be a limiting

factor in studying human populations for several years to come, and the shortages may be so severe that the initial choice of human population study centers for support should be dictated by the adequacy of their training facilities. Some human population study centers, therefore, should be sponsored by universities or have academic affiliations. When the more pressing manpower problems are resolved, efforts should then be directed toward establishing additional representation of both community-based and the more specialized population studies, widely dispersed throughout the United States and covering a broad spectrum of ethnic groups and general environmental and occupational exposures.

There will be a need for a central study group, functioning as a professional service unit, to keep track of new developments in human population study center activities, to conceive and stimulate population studies, and to view the programs in the perspective of broad national interests in a manner which will not conflict with specialized local study interests. Without interchange of information on the characteristics of the population covered and the types of data collected and collated, no intelligent matching of study objective and design with a suitable human population study center can take place. Interchange of information on methodological techniques would be particularly valuable in view of the fluid state of the art.

In view of the expense of delineating and setting up a human population study center, investigations restricted to cancer or to any single disease may be regarded as inefficient. Perhaps a more important consideration, however, is that restriction of studies to a single disease would restrict investigations of competitive morbidity and mortality risks of various diseases and the interplay of their effects on the population at risk.

National Cancer Institute Program

A National Cancer Institute intramural program concerned with human population study centers must consider the history of work undertaken by its staff as well as take note of human population study centers organized and

sponsored by other institutions. Work now in progress at the National Cancer Institute has generated a number of human population study centers, potential or in being. These include:

- Populations in Hagerstown and Washington County, Md., where major effort has centered on measurement of general environmental background effects.
- Population registers for the Papago and Pima Indian tribes, leading to the accumulation of a variety of morbidity and mortality observations on Indians of the southwest as they undergo the transition from reservation to urban life.
- A cohort of 800,000 persons examined in a mass chest X-ray survey in Pittsburgh, Pa.
- Employee populations of industrial plants in Pittsburgh, Pa. The possibility exists of defining large employee populations under continued medical surveillance in other localities as well.
- A cohort of railroad employees whose retirement and other benefits are administered by the Railroad Retirement Board.
- Veterans holding Government life insurance policies, from whom smoking histories have been obtained, who remain available for collection of histories on other items and who might be sampled for a variety of clinical and laboratory observations.

To these centers may be added such other possibilities as establishment of human population study centers in Connecticut and rural Iowa. These proposals have arisen as possible extensions of cancer registers and morbidity surveys for further studies in depth of these populations. Most of these human population study centers are in an embryonic state of development. At the moment they are run on a very modest scale, and by cumulation of data over several years yield information on morbidity and mortality risks among the population subgroups covered.

Intramural programs should, of course, meet the same high standard of staffing and planning expected from applicants for research grants. Proposals to proceed to more elaborate data collection in any of these potential human population study centers should have definite objectives, and data on precursive states should be a prime consideration. The interdisciplinary

staff requirements would be the same as for human population study centers under other auspices; any human population study center in the intramural program should incorporate the joint contributions of the Biometry, Carcinogenesis Studies, Diagnostic Research, and Epidemiology Branches of the National Cancer Institute.

Certain lines of inquiry may be particularly appropriate for the National Cancer Institute to undertake, such as those involving collaboration with investigators in other countries or with governmental agencies. One example is the study of cancer risks among the foreign-born and their immediate descendants to be undertaken as part of a parallel investigation of cardiorespiratory diseases. Access to the 1960 census data has been obtained in order to define households for which the head is foreign-born. Studies of these families will involve collaboration with investigators in their countries of origin, including England, Norway, Sweden, and Poland. Other populations abroad are of interest in their own right because of unusual features in the site distribution of cancer risks. Promising human population study centers could be established in such countries as Finland and South Africa.

Another illustration of intragovernmental agency collaboration would be work with the Veterans Administration and the National Research Council Follow-up Agency on a series of longitudinal studies of subgroups of World War I and II veterans with defined characteristics, with emphasis on the association of cancer risks with other disease entities.

Further suggestions for National Cancer Institute intramural activities might place a premium on projects beyond the resources of any other single institution. This might mean that the National Cancer Institute or the National Institutes of Health assume an active role in the

collation of findings from a network of human population study centers, as well as concerning itself with specific study enterprises. In advancing these suggestions, we would reiterate that the present powers of maneuver are limited by available personnel; recruitment and training of new staff must be pursued aggressively, in order for the National Cancer Institute to implement even a modest fraction of the opportunities for study of human populations.

REFERENCES

- (1) Turner, V.: Hagerstown Health Studies. An annotated bibliography. Public Health Bibliography Series No. 6. PHS Publication No. 148. U.S. Government Printing Office, Washington, D.C., 1952.
- (2) Dawber, T. R., Meadors, G. F., and Moore, F. E.: Epidemiologic approaches to heart disease: the Framingham study. *Am. J. Pub. Health* 41: 279-286 (1951).
- (3) Dorn, H. F.: Tobacco consumption and mortality from cancer and other diseases. *Pub. Health Rep.* 74: 581-593 (1959).
- (4) Health studies of human populations. Conference summary. *Pub. Health Rep.* 76: 955-965 (1961).
- (5) Kodlin, D., and Thompson, D. J.: An appraisal of the longitudinal approach to studies in growth and development. Monographs of the Society for Research in Child Development, vol. 23, No. 67. Child Development Publications, Purdue University, Lafayette, Ind., 1958.
- (6) Neel, J. V., and Schull, W. J.: The effect of exposure to the atomic bombs on pregnancy termination in Hiroshima and Nagasaki. National Academy of Sciences-National Research Council No. 461. Washington, D.C., 1956.
- (7) Hollingsworth, J. W.: Delayed radiation effects in survivors of the atomic bombings. *New England J. Med.* 263: 481-487 (1960).
- (8) Morris, J. N., Heady, J. A., and Barley, R. G.: Coronary heart disease in medical practitioners. *Brit. M.J.* 1: 503-520 (1952).
- (9) Morris, J. N., Heady, J. A., Raffle, P. A. B., Roberts, C. G., and Parks, J. W.: Coronary heart-disease and physical activity of work. *Lancet* 2: 1053-1057 and 1111-1120 (1953).