

#### IV. STATISTICS AVAILABLE FOR THESIS TOPICS IN A HEALTH DEPARTMENT

CARL L. ERHARDT

Information about individuals and statistical data regarding aggregates are available in a health department on a comprehensive, community-wide basis, but usually for vital events only. Vital statistics, therefore, cover the entire population, but routine information regarding the persons to whom the records refer is rather meager for intensive sociological research.

Services of a wide variety are rendered to individuals by a health department. These range from prenatal care to identification of persons with specific infectious illnesses and their treatment and follow-up to similar identification of persons with chronic diseases of later life and their referral, usually to other sources for care. These services, however, are restricted, either by design or otherwise, to selected population groups. Generally these populations are drawn from the lower end of the socio-economic scale, but even when they are not, they are often a self-selected group. Consequently, an investigator may be unable to generalize to any larger population. Nevertheless, useful studies of groups who resort to a health department for specific services are possible.

As with vital statistics, however, social information regarding the people receiving services is scant, usually limited to demographic descriptives. In addition, medical data will be found, but even such data will vary considerably depending on the type of service. With

tuberculosis patients, for example, who may be under long-term supervision, extensive medical histories may result. On the other hand, a patient volunteering for a screening examination may be asked to provide little in the way of medical facts, while the particular screening facility may add the results of only one or two specific tests, such as those for diabetes and/or glaucoma and/or a chest x-ray.

There are, in some areas, clinics intended for general workups of well persons, presumably on a continuing (perhaps annual) basis. These clinics will have a wide variety of medical information, such as blood chemistry, electrocardiographic tracings, proctoscopy, gastro-intestinal series, if indicated, as well as serology, tenometry, and chest x-rays. Even in such clinics, however, social information regarding the patients would be meager. Occupation may be given; in areas where a means test is applied, family income may be included. Information on family structure may be available on patients' charts, where continuing service is anticipated, but any service of a one-time nature would exclude such details.

Hence statistical data already compiled or readily accessible, that would be useful for sociological studies of persons served by a public health agency are rare. Of what use, then, can health department records be in sociological research?

#### AREAS AND METHODS FOR SOCIOLOGICAL RESEARCH

Ingenuity sometimes produces what is not apparently available. For example, the influence of religion (in addition to other factors) on marriage survival rates has been investigated in much detail by Burchinal and Chancellor in Iowa. Such studies were made possible by the simple expedient of matching divorce records in the state with the corresponding marriage records.<sup>1</sup>

Similarly, matching of records on an individual basis has allowed the analysis of infant mortality in New York City in terms of variables reported only on the birth certificate of the child, such as age of mother, parity, residence, or weight of baby at birth and type of delivery, or residence as a socio-economic variable.

Such analyses must necessarily be limited to factors available on the documents matched but, as mentioned before, existing health records give little information of specific sociological import. It seems to me, therefore, that in most instances, health records can be viewed best as sampling frames from which groups may be selected for collection of further information that will be useful in establishing sociological determinants. For example, the Division of Vital Statistics of the National Center for Health Statistics is collecting for samples of live births and infant deaths such items as prenatal care, family income, and occupation and working history of the mother in order to assess the influence of these factors on infant loss.<sup>2</sup>

Another study proposal presented to the New York City Department of Health aimed at ascertaining factors in the treatment environment (philosophy, policies, and practices) as well as social characteristics of the patients that influenced persons with tuberculosis to accept or reject care.<sup>3</sup> This question is one of a rather general class that is of major consequence to a department of health. What are the social, economic, and cultural patterns that determine who will take advantage of a service offered? How do these patients differ with respect to such characteristics from those who ignore or refuse the service? The latter question, of course, involves obtaining like information for a sample of those who have not availed themselves of the service. Selection of the sample of those accepting service may be relatively easy. How to define the population from which this group is drawn and how to select the sample of nonusers may be exceedingly difficult. But a researcher will inevitably have to meet such problems in his professional career. It can be a salutary experience in preparing for his doctorate to learn that life has to be met as it exists, not as theory demands.

A recent report in *The American Journal of Public Health*, by Horn and Waingrow,<sup>4</sup> shows that health education efforts (probably among other factors) have been successful in improving knowledge of cancer and its symptoms. Not only is such knowledge shown to have increased, but the increase has occurred among cohorts of different birth years and sex. Another aspect of importance is that the message regarding the need for regular examinations is being

heeded, as evidenced by the substantially higher proportions in 1962 than in 1940 of persons who have had complete physical examinations within two years of interview. This study is an example of collection of current data to measure the extent of change since a base period that may have resulted from programs introduced and the degree to which population groups have been differentially affected, if at all. The question still remaining is identification and study of those groups which resist change for guidance in learning how to be even more effective.

#### NEW DIMENSIONS NEEDED

My last comment leads me to express a somewhat troubled feeling about the impression I receive from the title of this panel discussion. A doctoral dissertation, in my understanding, reports research that adds to knowledge or understanding. Therefore, the candidate for the doctoral degree cannot satisfy the requirements, it seems to me, by analyzing data that someone else has collected and tabulated. Some new dimension must be added that yields new information or new interpretations.

Sometimes data have been collected but not tabulated or analyzed fully. Retabulation of such data to determine the effects of one or more previously neglected variables may be a real contribution. However, I question whether such a procedure gives the doctoral candidate the requisite training and experience needed later to conduct full-scale research. Major aspects of the research process, once a question or hypothesis has been appropriately framed, include establishment of definitions of terms, determination of the best method of collecting the required data, designing the collection instrument, determining the best phraseology for questions, testing the instrument, devising codes if necessary, and planning tabulating procedures and tables. The use of existing data removes all these processes from the activities of the candidate and deprives—I use the word deliberately—him of the highly instructive labor that must go into a research project before any data are collected.

I do not mean that existing material should be ignored in seeking an answer to a pertinent question. Obviously, it is senseless to repeat a collection process when satisfactory data have already been assembled. Projects to exploit such data are quite proper research efforts. My question relates to whether this is an appropriate stage for a doctoral candidate to start his research. Analysis and the drawing of inferences and conclusions represent the most satisfying part of research. Limitation of the candidate's dissertation to analysis may result in a belief that this aspect of research is his sole responsibility. Otherwise how does he learn about the bitter, disappointing experiences that may be met and how to surmount them? How does he learn that interpretation of the data involves an intimate knowledge of what went into them? How does he learn that close supervision at every point of the way is as important as, if not more important than, high-powered analytic techniques and that such techniques cannot overcome basic deficiencies in the observations?

One young doctor, appointed to set up the research aspects of one of our new programs, came to visit me shortly after his appointment. Not a word was said about what knowledge was needed for direction of the program, what sources might be tapped, what questions were crying for answers. His sole concern was the configuration and capacity of our computer center, since he was going to need access to it constantly. And this was before he had any idea of what his research was to involve! His obvious consternation when I told him our poor benighted department was still struggling with EAM equipment may have been the cause of his resignation a few weeks later. I'm not quite sure. However, I am sure that it was assumed by the department, because he had a doctorate, that he was adequately prepared, by training *and experience*, to pursue independent research starting from scratch.

## SUMMARY

In summary, my message is that data about people are available in a public health department in connection with a wide variety of programs affecting persons of different ages. Information regarding such persons is usually limited to the common demographic descriptives plus some limited medical facts determined by the nature of the program. (Where special research efforts are being made, additional data may be available but this would not be the common situation.) Tabulations reflecting demographic variables are often regularly prepared. They have usually been fairly well exploited, and new knowledge or enlightenment can come only from collection of additional basic data.

Any health department, however, has many questions regarding the people to whom its services are proffered. The answers to these questions might permit it to give a better service or extension of the service. To obtain the answers will require perspicacity, ingenuity, and much sheer determination as well as technical knowledge. But is it not an objective of the doctoral program to probe for and develop these qualities in the candidates?

## REFERENCES

<sup>1</sup> Burchinal, Lee G., and Chancellor, Loren E., Survival Rates Among Religiously Homogamous and Inter-religious Marriages, *Research Bulletin No. 512*, Agricultural and Home Economics Experiment Station, Iowa State University, December, 1962.

<sup>2</sup> Sirken, Monroe G., Communication to Commissioner George James, New York City Department of Health, February 19, 1964.

<sup>3</sup> Taylor, Ruth B., et al., Project outline prepared for the New York City Health Department Research Committee, September 30, 1958.

<sup>4</sup> Horn, Daniel, and Waingrow, Selwyn, What Changes Are Occurring in Public Opinion Toward Cancer: National Public Opinion Survey. *The American Journal of Public Health*, 54, 431-440, March, 1964.