INTERNATIONAL COMPARISONS
OF HEALTH SERVICES SYSTEMS

KERR L. WHITE

Social arithmetic, like society itself, evolves in purposeful ways. First one counts people—the population census. Next one counts food production—agricultural statistics. Then come the counting of jobs, industrial productivity, income and expenditures—labor and economic statistics. That is followed by the counting of deaths, and later of disease, disability, discomfort, deliquency and disruption—health and social statistics. Finally, one tries to measure the quality of the environment and of life itself—the use and enjoyment of books, poetry, music, drama, art and recreation, and finally the growth and fulfillment of individual lives—the "human" census. Statistics, Bradford Hill used to say, are really people with the tears wiped off.

A country's stage of development is the major factor influencing the development of its statistical systems. But why are statistics needed? Empirical observation, intuition, guesses, instinct and impulse are the customary ways of making most decisions, most of the time, for most people. When societies organize their talents and agree on certain aspirations or objectives, social arithmetic constitutes the only objective way of describing events and documenting experience, and perhaps the most important way of encouraging rational decisions about major problems facing large groups in society. Most countries conduct a periodic census of the population; no country conducts a "human" census. Many produce agricultural, labor and economic statistics and a number produce health statistics that vary in their comprehensiveness and sophistication.

117

Lacking a universal prescription for Utopia, and recognizing that most human enterprises are run by human beings—with all that implies for human affairs—the search continues for ways of improving decision-making processes. The computer both removes the burden of routine and trivial decision-making and promotes forms of cooperation to which most mortals are quite unaccustomed. Aspirations, objectives, ends and alternative means and resources needed to achieve these can be compared by computers, but the fundamental decisions and choices remain human decisions and choices. The computer also frees human time and energy to devote to these critical decisions about human ends and values.

These matters are discussed at some length to establish the place of health statistics in the context of social science and health planning. Planning deals with decision-making as it affects ends and means; the social sciences deal with human behavior and its effects on ends and means. Health statistics reflect one outcome of human behavior—health, and the several outcomes of the health services sanctioned and supported by society to maintain and promote the health of its individual members. The central issues in health planning, as in other aspects of social planning, concern the establishment of priorities with respect to competing objectives and the allocation of scarce resources among competing priorities. It is apt to boil down to decisions that will make optimal use of health resources in providing effective and efficient health services.

To assist in the achievement of these objectives, it is appropriate to compare different systems for providing health services. It seems unlikely and undesirable for one optimal system to exist for providing health services suitable for all people at all times. The need, therefore, is to establish traditions, institutions and instruments for continuously evaluating and comparing health services systems and their impact on health. Such comparisons can be useful, if they are not essential, both for assessing the relative merits of different systems, and for understanding those cultural and social forces that condition the transactions between the providers and consumers of health services and knowledge.

Regional comparisons within countries may ultimately be as useful as comparisons between countries. At present, however, denominator data for regions within many countries bearing on characteristics of the populations, on the health resources available to these populations and on the outcomes of health services provided, are not readily

available. An even greater limitation is lack of experience with methods available to obtain local and regional data.

A real need, therefore, is to undertake international comparisons of health services, using compatible, if not comparable, definitions, methods and questions to understand differences and similarities in health services systems, and to understand the principles of human behavior that condition the medical care process in general and specific preventive and therapeutic transactions. A corollary of the emphasis on the use of health statistics to understand health services and health behavior is the need to distinguish the impact of health services from the impacts of improvements in education, the physical and social environment, material productivity, communications, transportation, economic security and the uses of leisure. Planners need to be able to measure the size of health problems, identify their locations and assess the specific impact of health services on them.

All of this is easy to say, and may sound unduly idealistic, but in fact much has been done already. Health statistics have been used over many years for administrative purposes, but no country, as yet, has a comprehensive health statistics system. Most emphasis has been placed on vital statistics, and particularly on death statistics. Modest emphasis has been placed on measurement of disease, through sickness surveys and epidemiological studies of specific diseases. Hospital data frequently are limited to "activity" data describing admissions or discharges without associated demographic or diagnostic information. Information about illness seen in physicians' offices and clinics and the services physicians render is almost totally lacking. The ideal health services information system for the population of a community, region or country should include demographic data about individuals, physicians and hospitals, and diagnostic information describing the illnesses precipitating demand for services, and the nature of concurrent or complicating conditions. It should be possible to associate such events as death, admission to a hospital or related institution, a visit to a physician or other health professional, with the prevalence of perceived morbidity, professionally defined and scientifically confirmed morbidity of similar or related symptoms and conditions in the population.

The objectives of the health services system of society are the reduction of premature death, disease, disability, discomfort, delinquency and disruption—six "D's".[1] To accomplish the first two, specific or curative measures of documented efficacy are needed; i.e., medical

"cures." To accomplish the other four, a variety of palliative, supportive and informative maneuvers and instructions are required. Perhaps only about ten to 20 per cent of the procedures employed by health professionals have objective, controlled clinical trials to support the view that they are likely to be more beneficial than harmful or useless for the purposes for which they are used. About 20 to 40 per cent of most therapeutic improvement seems to be associated with either the placebo or Hawthorne effects. Considerable uncertainty surrounds the other 40 to 70 per cent of activities that occupy the time and effort of health professionals. Virtually nothing is known about the nature of benefits derived from them. One of the major objectives of a health services information system would be to increase the application of efficacious procedures or treatments to all those who could benefit from them; that is, to increase the effectiveness of health services. The results might be reflected in the reduction of premature death and disease. A second major objective would be the reduction of disability, discomfort, delinquency and disruption. A third major objective would be improvement in the efficiency or economy with which stated objectives are achieved.

From these equations, any discussion of health manpower, health facilities and health activities has been omitted. These are means, important means, but only means to the achievement of objectives. It is the establishment of clear, reasonable objectives within countries that facilitates comparison of different health services systems designed to achieve stated objectives. Certainly, an account of activities—patients seen, patients hospitalized, visits made, pamphlets distributed, injections given and x-rays taken—is of limited value with respect to the kinds of objectives set forth. This may be interesting and indeed essential preliminary information, but it reports "activity" not "achievement." "Being busy" and "doing good" no longer seem adequate justification for the vast budgets requested by the national health services establishments of the contemporary world. However, before "activity" or utilization patterns can be modified in desired directions, it is essential to identify those independent variables that influence them.

How can disease, disability, discomfort, delinquency and disruption be measured in comparable fashion in different countries? With respect to delinquency and disruption, little can be said. In the full spectrum of human distress, however, delinquent behavior whether from school, work or other accepted responsibilities, is increasingly regarded as a form of illness requiring educational or theraupetic,

rather than punitive intervention. Similarly, social disruption as reflected in broken homes, ineffectual life patterns, and personal dissatisfaction, is recognized as an integral behavioral component of health. They are difficult to define and more difficult to measure and, apart from stressing the need for their inclusion and for methodological studies bearing on them, they will not be discussed further.

Comparisons, whether between or within countries or regions, are possible only when comparable definitions and methods are used, and when standardized rates with standard errors of the estimates are provided. Comparisons of crude rates without standard errors collected by different methods employing different definitions are apt to be meaningless, if not misleading. National health and sickness surveys, have been widely used[2,3] for national purposes. They may even permit gross estimates of major problems[4] and they undoubtedly represent essential precursors of international comparative health surveys.[5]

What are some of the problems encountered in undertaking such international surveys? First, it is important to limit the objectives of the survey. A health examination survey is quite different from a health interview survey and both are different from a health record survey. Each has its appropriate role to play in the generation of health statistics. To attempt international comparisons using all three simultaneously is probably unwise, given the current availability of manpower, skills, experience and resources. The problems seem least formidable for health interview surveys based on probability samples of households. Health record surveys are limited by wide variability in the ways that national, institutional and professional record systems are structured and accessible; a great deal of methodological cooperation will be required before comparability can be achieved. Health examinations, particularly certain screening examinations, may be feasible for international comparisons, provided relatively simple tests of known sensitivity and specificity for detecting specific diseases or physiological aberrations are available.[6] The problems of controlling standards are formidable but not impossible, and the problems of observer variation are probably as great, if not greater, than those encountered in household interviews, if only because physicians and nurses are less accustomed to the idea of observer error and variation than are social scientists.

The health interview is a feasible method of undertaking international comparisons of illness and health. Five major kinds of problems must be considered, which are not mutually exclusive.

121

1. *Direct measurement of morbidity.* It is difficult to undertake a morbidity survey, a utilization survey and an attitude survey simultaneously. Emphasis, flow of questions, length of questionnaire, sample design and size probably preclude that. Since the "hardest" data obtainable by this method, and the most important variable to understand first appears to be "utilization" it seems best to make this the point of departure in the interview and the major dependent variable to tackle first, bearing in mind that "utilization" measures "activity" and not "achievement." Morbidity may be divided into three levels: a. Lay perceived morbidity; i.e. the morbidity perceived by individuals, their families or friends, and hence the morbidity that is a primary determinant of utilization; b. Professionally perceived morbidity; i.e. morbidity perceived initially by a nurse or primary physician at the time of the patient's initial visit or by an epidemiological questionnaire, the validity of which has been established previously; c. Scientifically confirmed morbidity; i.e. morbidity determined on the basis of laboratory tests, x-rays and clinical findings.

If the emphasis of the survey is on utilization of health services, then the most relevant morbidity component is lay perceived morbidity. This level of morbidity can be expressed in terms of problems, symptoms and conditions, using colloquial language; many of the frequencies will be high and the associations with utilization relevant for both respondents and lay interviewers. Whether the responses are scientifically "accurate" is not a problem, because initiation of the medical care process is much more a function of the patient's perceptions and behavior than of the physician's.

2. *Medical coding.* Almost half the problems presented initially to primary physicians cannot be coded readily in accordance with the International Classification of Diseases.[7] Sample frequencies for many diseases are too low to permit any analysis with respect to utilization or even demographic characteristics. Pre-coded lists of selected, common, acute and chronic symptom/conditions may therefore represent a preferable approach to estimating general morbidity in populations. If medical conditions reported by lay respondents to lay interviewers are to be coded by lay coders, the possibilities of bias and error are substantial and careful controls to estimate variations between coders in the same and in different countries would be essential. This problem is hard enough in the

coding of death certificates and hospital discharge diagnoses to post-pone its application to household survey data for some time.

3. *Basic definitions*. International comparisons require the use of identical definitions. Among these are: doctor, nurse, dentist, hospital admission and discharge, household and respondent. Not a few problems are encountered in developing definitions that will produce comparable information from different countries with differing health services systems. The words themselves and their literal translations are of less importance than the comparability of the information elicited in response to the question.

4. *Social and cultural variables*. The problems encountered in measuring expectations, attitudes and satisfactions bearing on the use and receipt of health services seem to present fewer problems in international comparisons than do the other issues discussed. Questions that appear to scale satisfactorily can be used to measure "tendency to use medical care services," "skepticism toward medical care" and "health orientation."[8] These in turn can be related to utilization data. On the other hand, the classification of education, occupation, social position and income present major problems. Both occupation and income data are relatively difficult to obtain and classify in any household survey, and international comparisons compound the problems. It should be possible, however, to divide the full range of data bearing on these four parameters into thirds or quartiles, each with large enough numbers to permit meaningful comparisons between the rates for similar thirds or quartiles in different countries.

5. *Communication*. These and many other problems require discussion in detail by all participating investigators in a collaborative study. Working conferences, task force meetings, use of consultants, extensive travel, cables and international telephone calls are all required.[9, 10] Much more important than formal communication is the informal communication that assures the levels of understanding, agreement and commitment required to produce truly comparable data. Without substantial human and fiscal resources to support adequate communication between collaborating regions and countries, an exercise in international comparisons of health services is unlikely to succeed.

To summarize, it does seem possible to measure independent vari-

ables such as disease expressed as lay perceived morbidity (e.g., problems or symptom/condition complexes), certain types of specific morbidity (e.g., visual impairments), disability (e.g., activity limitation, bed-disability), discomfort (e.g., pain or distress) and attitudes toward health care and health services, and to relate these to dependent variables such as drug consumption, visits to nurses and doctors and admissions to hospitals. Controlling factors such as age, sex and marital status can be used along with such factors as education, occupational class, social position and income. Data of these kinds can be expressed as standardized rates, together with standard errors of the rates.[5]

The next phase is to apply these methods in understanding observed variations in utilization, and to relate these in turn to outcome data and "achievement" variables reflected in rates for disease, disability and discomfort. Multivariate analysis can be used to determine the relative contributions of independent variables to variations in utilization. Some of the residual variations may be functions of the health manpower and facilities available to the populations studied, some to differences in the health services systems, including methods of financing, staffing and organizing health services, and some to the original independent variables. It is a circular system, and all the parts cannot be studied simultaneously with equal precision.

It would be illusory to think that the stage has been reached where direct comparisons of the elements comprising health services systems can be pursued internationally with any degree of confidence. What is needed is intensive work, improved communication and the development of traditions of collaboration between social scientists, epidemiologists, clinicians, statisticians and administrators in different countries so that health services in many lands may be continuously compared and improved.

REFERENCES

[1] Sanazaro, P., Seminar on Research in Patient Care, *Medical Care*, 4, 43–50, 1966.

[2] National Center for Health Statistics, *Origin, Program and Operation of the U.S. National Health Survey*, Series 1, No. 1, Washington, D.C., United States Department of Health, Education and Welfare, 1963.

[3] World Health Organization, Continuous Health Surveys of the Population, EURO 215(3)/5, 2 February, 1967 (Processed), Geneva, World Health Organization, 1967.

⁴ White, K. L., Greenberg, B. G. and Williams, T. F., The Ecology of Medical Care, *New England Journal of Medicine,* 265, 855–892, 1961.

⁵ White, K. L., *et al.,* International Comparisons of Medical Care Utilization, *New England Journal of Medicine,* 277, 516–522, 1967.

⁶ Wilson, J. M. G. and Jungner, G., *The Principles and Practice of Screening for Disease* (Processed), Geneva, World Health Organization, 1967.

⁷ Research Committee, College of General Practitioners, The Records and Statistical Unit, *Journal of the College of General Practitioners,* 6, 195–204, 1963.

⁸ Bice, T. W., personal communication.

⁹ Committee for International Collaborative Study of Medical Care Utilization, *The Chronicle of a Feasibility Study: 1964–66,* International Collaborative Study of Medical Care Utilization (Revised), 1966.

¹⁰ Mabry, J. H., *et al.,* The Natural History of an International Collaborative Study of Medical Care Utilization, *Social Sciences Information,* 5, 37–55, 1967.