

THE CONCEPTUALIZATION AND MEASUREMENT OF HEALTH FOR
POLICY RELEVANT RESEARCH IN MEDICAL CARE DELIVERY

John E. Ware, Jr.

February 1976

P-5599

The Rand Paper Series

Papers are issued by The Rand Corporation as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of Rand's contracts or grants. Views expressed in a Paper are the author's own, and are not necessarily shared by Rand or its research sponsors.

The Rand Corporation
Santa Monica, California 90406

ACKNOWLEDGEMENTS

To a large extent, this paper represents the contributions of a number of individuals at the Rand Corporation who have been working on problems of health status assessment. While assuming full responsibility for the summary of our work as presented in this paper, I would like to gratefully acknowledge the contributions of: Kenneth H. Applegate, Robert H. Brook, George A. Goldberg, Lisa A. Hahn, Kent H. Marquis, Joseph P. Newhouse, Anita L. Stewart, Laura L. Tosi, and Kathleen N. Williams.

ABSTRACT

This paper discusses several issues regarding health status assessment for purposes of evaluation of medical care delivery. The issues include: 1) reasons for health status assessment, 2) the nature and number of health concepts that can be measured, and 3) some of the implications of various measurement strategies. The Health Insurance Study (HIS), which The Rand Corporation is conducting for the Department of Health, Education and Welfare, is offered as an example of a social experiment in which the measurement of health will aid in policy decisions about how medical care should be delivered. Solutions to the problems of health status assessment for purposes of the HIS are offered as examples that may have general applicability. It is argued that: 1) the use of health care services, which in the past has been treated as an outcome measure, is not sufficient, i.e., more use doesn't imply better care; 2) an omnibus approach to the conceptualization and measurement of health status is required; 3) differences between disadvantaged and nondisadvantaged groups with respect to the reliability and validity of scores computed from survey measures of health must be kept in mind; and 4) self-ratings of health should be given greater emphasis in evaluating medical care.

THE CONCEPTUALIZATION AND MEASUREMENT OF HEALTH FOR POLICY RELEVANT
RESEARCH IN MEDICAL CARE DELIVERY*

John E. Ware, Jr.**

Health, defined as physical and mental well-being, appears to be the most salient human value in America today.^{1,2} When asked to rank the importance of 18 end states or guiding principles of life, adults in five independent studies ranked health first more often than each of 17 other values including National Security, Salvation, Freedom and Love.³ The striking salience of health as a value in relation to other human values suggests that health needs are not being satisfactorily met and that health problems constitute an important constraint on the quality of American life. Consistent with this interpretation is the finding that the health care disadvantaged (i.e., the less healthy and those who are less able to gain access to health care services when they are needed) tend to value health most highly. The health care disadvantaged also tend to be the disadvantaged as defined by socioeconomic criteria; i.e., they are less educated, poorer and older.

Given the salience of health, it is not surprising that the financing of health care is among the most critical policy issues in America today. One way to evaluate the effects of policy options regarding health care delivery and financing would be to compare options in terms of their effects on use of health care services. However,

*This paper was prepared for presentation at the annual meeting of the American Association for the Advancement of Science in Boston, Massachusetts, February 18 to 24, 1976. The research reported herein was performed pursuant to the Health Insurance Study, a grant from the U.S. Department of Health, Education and Welfare, Washington, D.C. The opinions and conclusions expressed herein are solely those of the author, and should not be construed as representing the opinions or policy of any agency of the United States Government.

** Dr. Ware is a research psychologist on the staff of the Health Insurance Study and in the Social Science Department, The Rand Corporation, Santa Monica, California, 90406.

estimates of differences in use of health care services as a function of differences in cost-sharing would not, alone, be satisfactory for purposes of evaluating policy options. Rather, direct estimates of effects of differences in cost-sharing on health status are required. Why is this so? If the relationship between use of health care services and health status outcomes was strong, then measures of use of health care services might provide an adequate approximation of effects of cost sharing on health. However, knowledge of the quantity of use of health care services provides little insight into the quality of care given. Further, care of optimal quality and quantity has not been adequately linked to health status outcomes. Thus, direct measures of health status outcomes are required in order to evaluate policy options regarding health care financing.

THE CONCEPTUALIZATION OF HEALTH

What is health? As we have noted above, health is one of the most sought after end states or guiding principles of life. Health is also a state of being with extremes of perfect health and death. Perfect health can be considered as the ideal or optimum state in which every cell and organ as well as the organism as a whole is functioning perfectly.

Attempts have been made to characterize optimal health in terms of distinguishable health components, the absence of illness and the achievement of positive well-being. For example, consider the conceptualizations offered by the World Health Organization (WHO) and the American Public Health Association (APHA). The WHO has defined three dimensions or components of health including physical, psychological, and social well-being. The WHO definition implies that optimal health is a state characterized not only by the absence of physical, psychological, and social limitations but also by achievement of a positive state with respect to all of these health components. It has also been suggested that physiologic health, which denotes the state of individual organs

of the body, should either be added into the concept of physical health or considered a fourth dimension in addition to those defined by the WHO.⁴

A number of measures which appear to define one or more of the major components of health, as defined by the WHO, have been published, including measures of physical health,⁵⁻⁸ psychological health,⁹⁻²² and social health.^{11,22-26} Measures of functional limitations due to health seem to be tapping both physical and social components of health.²⁷⁻³² A very comprehensive approach to the quantification of physical, psychological and social function and dysfunction is that being undertaken using the Sickness Impact Profile.³³⁻³⁵ Although good measures of more than one of the major health components are infrequently included in the same study, some information about the relationships among measures of health components is available. A positive relation between measures of psychological and social components of health has been reported²³⁻²⁵ and a positive relationship between measures of physical functioning and psychological components of health has been reported.¹⁹ Furthermore, there is support for the conclusion that three components of health can be distinguished even when the same method of measurement is used. Factor analyses of correlations among ratings of the importance of health states yield correlated factors corresponding to physical, psychological, and social components of health.³⁶ The positive correlations among measures of health status components, when measures of more than one component are included in the same study, are consistent with the notion of a general underlying health factor.

The American Public Health Association (APHA) has defined four levels of health which can also be viewed as a hierarchy of the goals of medical care. The goals at the four levels are: (1) conservation of life, (2) prevention, control and treatment of serious morbidity, (3) prevention, control and treatment of minor morbidity (illnesses

that cause inconvenience, annoyance, economic loss) and (4) attainment of positive well-being (vigor, happiness, etc.).

The APHA levels and goals appear to be optimal and not necessarily consistent with current medical care practices. Most medical care activity in America has been directed toward the first two APHA goals, namely, conservation of life and the prevention of serious morbidity. Medical care in the future may be directed more toward goals three and four, namely, the prevention and treatment of minor morbidity and promotion of positive well-being. The treatment of all serious morbidity may not be cost effective in the future. Rather, it may be that the total amount of suffering in America, due to poor health, will be reduced more vis-a-vis prevention and treatment of minor morbidity and promotion of positive well-being as opposed to treatment of all serious morbidity.

ALLOCATION OF MEASUREMENT RESOURCES

Given that physiologic and physical, psychological and social components of health should be measured, it does not necessarily follow that measurement resources should be evenly divided among the three components. Other considerations are involved in developing a total measurement strategy, including the extent to which: 1) medical care can change or affect each component, 2) variability in scores for each component occurs in the general population, and 3) the relative payoff from changing each component is the same. With respect to the third consideration, the assumption of equal payoffs is plausible.

Effects of Medical Care

One primary question in determining the emphasis of measurement is the degree to which the component(s) of health selected for measurement are likely to be affected by the medical care system. Attainment of positive well-being (stage four as defined by APHA) and high scores for

psychological and social components of health (as defined by WHO) involve phenomena of society beyond the control of medicine (e.g., transportation, education, entertainment, and so on). Clearly medical care, as practiced in America today, impacts most on physical health and to a lesser degree on psychological and social health; therefore, greater emphasis should be placed on measurement of physical health.

Interest in positive well-being and the psychological and social components of health becomes only secondary. The goal of measurement in the latter two areas is to better understand what contributes to positive well-being and the relationships among scores for the health components. Even though social and psychological aspects of well-being may be given little attention by the formal medical care system, some increase in their understanding may be helpful in progressing from a "quantity" to a "quality" of life concept.

The same logic applies to the allocation of measurement resources within a major component of health. For example, within the psychological components of health, conditions or diagnoses thought to be most susceptible to intervention by the formal medical care system should be emphasized. Specifically, resources should be allocated first to the measurement of depression, anxiety, and disruption of self-control over thought and behavior.

Finally, it should be noted that not all effects on health can be detected in any one experiment. For example, constraints on sample size and study duration are likely to preclude a look at the effects of differences in medical care on mortality. On the other hand, measurement of changes in serious and minor morbidity can be accomplished in a large scale social experiment.

Variability in Health Scores

A second consideration in the allocation of measurement resources is the amount and nature of variability in the distribution of scores for

each health measure. The distribution of health scores is, of course, largely a function of how health is defined and measured. Before discussing some implications of this phenomenon, let us clarify the extent to which score distributions differ. Consider, for example, the distributions of scores for scales to measure chronic functional limitations (constructed according to the criteria of Guttman scalogram analysis) and a scale to measure the perception of health (constructed according to Likert scaling criteria) in a sample of approximately 1200 adults. When health was defined as the absence of each of three kinds of functional limitations, approximately 90 to 98 percent of respondents enjoyed perfect health. A perfect score for each scale means that the respondent reported none of the chronic functional limitations defined by the scale items. In contrast, only seven percent of the same respondents had perfect scores on the scale to measure health status perceptions. A perfect score on the latter scale means that the respondent chose the most favorable rating on every item. These differences in score distributions have certain implications for health measures. First, conclusions about the proportion of persons who enjoy perfect health will vary greatly and perhaps erroneously as a function of the measurement technique employed. Second, the power of statistical techniques in a given study may vary greatly during hypothesis testing. Obviously, persons who have perfect scores at the beginning of a study cannot better their scores.

MEASUREMENT OF HEALTH

The preceding comments have been directed at conceptual issues relevant to policy research in medical care delivery. Arguments pertaining to why health should be measured, what (about health) should be measured, and the allocation of measurement resources have been offered. The comments below pertain to issues involved in how health should be measured.

There are many issues involved in deciding how health should be measured. The development of a total measurement strategy occurs

best within the context of the requirements of a specific investigation. Therefore, for purposes of this paper, solutions to the problems of health status assessment will be discussed in the context of a particular study, namely, the Health Insurance Study (HIS) which The Rand Corporation is conducting for the Department of Health, Education and Welfare. Let us begin with a brief summary of the HIS.

The Health Insurance Study

The HIS³⁷ is a large scale social experiment designed to assess how alternative insurance plans varying in cost-sharing arrangements, affect: 1) the demand for medical care services, 2) the health status of individuals, and 3) the quality of medical care. The experimental sample includes approximately 7500 individuals in 2200 families drawn from four areas of the country. Participants in the study are enrolled for a period of either three or five years. Low income families are being oversampled. Families are randomly assigned to 15 different insurance plans plus a control group. Eligibility for participation is quite broad. The only ineligible people are those 62 years of age and older, members of the military, and persons confined to institutions.

Omnibus Measurement Strategy

For reasons noted above (e.g., the need to measure more than one component of health and competing measurement strategies), an omnibus approach to the conceptualization and measurement of health in the HIS has been implemented. Adequacy of measurement has been taken, then, to be synonymous with comprehensiveness of measurement. Furthermore, because of the overlap between scores for components of health (physical, psychological, and social), a comprehensive battery is required to interpret experimental effects. This requirement holds true even when interest is limited to one of the major health components. Consider the following example. If more generous health insurance is observed

to produce a change in physical health, correlations between physical health measures and measures of other health components must be taken into account in order to be confident that the apparent effect on physical health is being properly interpreted. Therefore, regardless of whether one or three major components of health are of interest it seems that a comprehensive battery is necessary in order to control for the correlations among measures of health components when results are interpreted.

Sources of Health Information

Community health and health levels for specified populations can, in some instances, be estimated from available data regarding rates of mortality and serious morbidity. The validity of such data as an indicator of health has been questioned, and practical constraints limit the utility of health data that are part of the public record. For example, differences in scores for available indicators are certain to reflect the effects of many factors, only two of which are the structure and process of medical care.

Therefore, in order to place persons along the health continuum for purposes of experiments in medical care delivery, observations of health status by trained personnel and survey measures of health are required on an individual basis. Observations by trained observers are usually referred to as objective health measures and some survey measures of health are referred to as subjective health measures. Since there are many inconsistencies in the labeling of objective and subjective health measures, the various measurement strategies are best labeled in terms of how data are gathered, e.g., by self-report or independent observation, and without anticipation of their reliability, validity or even their objectivity.

A good source of information regarding physiological health is a multiphasic screening examination in which trained personnel use standardized measurement methods. Estimates of disability days can be

obtained, with reliance on very short recall periods, from a weekly or biweekly summary of a diary which is kept by the individual. Information regarding functional limitations, physical abilities, psychological symptoms, social participation and perceived health can be obtained from self-administered questionnaires, in most cases, with satisfactory results.

The Construction of Health Measures

Once all the issues discussed above have been considered, the method of operationalizing the health concepts must be selected. Wherever possible, available measures should be used, reflecting the "state of the art" of measurement which prevails at the time the research is conducted. Frequently, however, measures of sufficient reliability and validity are not available for all health concepts of interest. Given that only sparse resources are available for measurement research, the opportunity of a large social experiment should be used as much as possible to improve the reliability and validity of health measures. Such studies will not only insure that measures have met reliability and validity criteria for purposes of a particular investigation, they will also contribute to an improved understanding of measurement practices.

Recent innovations in health status assessment include the construction of standardized health scales from survey measures. These scales seem to represent improvements in the reliability and validity of health scores and the comparability of survey measures across studies. Use of health scales contrasts with the common procedure of computing a health score from the response to a single questionnaire item. Use of single-item health scores appears undesirable from the point of view of both reliability and validity,³⁸ single-item measures are less likely to achieve minimum standards of score reliability (particularly in studies involving disadvantaged persons) and, in certain instances, seem to be more highly correlated with estimates of response bias than

are balanced standardized scales. Examples of survey measures of health status constructed according to scaling criteria are the functional limitations measures described by Bush and others,²⁷⁻³² the Sickness Impact Profile,³³⁻³⁵ and scales to measure health status perceptions.^{31,38,39}

Existing measures of health status differ with regard to the component of health which is measured; some appear to primarily measure one component (physical, psychological, or social). Health scales also differ with respect to the method of placing people on the health continuum or into health categories. Some methods define health as the presence or absence of negative states; others represent attempts to include the extent to which positive health states are present. It is important to keep in mind that a score computed from a health status measure is likely to reflect both the particular health component being measured and method of measurement. Therefore, in addition to sampling measures of different health components, it is important to sample methods in a study like the HIS. This is particularly true with respect to the assessment of physical health, where a wide range of different methods are currently being advocated. Although many of these methods are being thoroughly studied, these studies are usually conducted in parallel. Thus, we know little about the relationships among the measures.

Reliability of Health Scores

The reliability of health scores, particularly those computed from survey measures of health status, does not appear to have been thoroughly studied. Reliability estimates based on our studies of the interrelationships among items within a scale indicate that scales to measure physical abilities, chronic functional limitations due to health and health status perceptions are quite reliable. Scores for functional limitations and health status perceptions also appear to be fairly stable over time. Findings regarding the stability of health scores suggest that analyses of data from repeated measures of health in a longitudinal experiment like the HIS are likely to improve the precision of hypothesis testing.

An issue that has not been adequately addressed is the extent and nature of differences in the reliability of health scores across subgroups of study populations. Reliability coefficients are only rarely computed separately for groups within total study samples. Our studies of the internal consistency reliability of health status perceptions scores indicate that scores are consistently and notably less reliable for groups of disadvantaged persons than for the nondisadvantaged.³⁸ The extent of these differences must be studied further for measures of other health concepts and for health scores derived from a variety of common scaling methods.

Validity of Health Scores

An important issue in the selection of a health measure is the validity of the scores it yields. Resulting scores should reflect differences in health status for persons manifesting a wide range of medical problems or diagnoses. The HIS provides an excellent opportunity to improve understanding of the validity of new and commonly used health status measures, by studying the interrelationships among a wide range of health status scales known to yield reliable scores. Two validity issues have already surfaced during preliminary analyses of HIS data. The first concerns the validity of health measures across diagnoses, i.e., Are health status scores equally valid for persons with different medical problems? The second concerns the validity of health scores across subgroups of study populations, i.e., Do health status scores have the same meaning for disadvantaged and nondisadvantaged persons?

In studies of cross sections of the general population, one should not rely entirely on health status measures that focus on specific medical problems or diagnoses. Preliminary findings indicate that measures of functional limitations, despite the fact that they appear to define generic categories of human functioning, may be more sensitive

to the presence or absence and severity of certain medical problems or diagnoses than for others.⁴⁰ Measures of health status perceptions, on the other hand, may reflect the presence or absence and severity of a wide range of diagnoses and problems (including problems of a psychological nature).

On the other hand, important differences in health status may not be detected if only generally applicable measures are relied upon. For example, for asthmatic patients, gross differences in severity of illness may be detected by general measures of health status perceptions, functional limitations, or physical abilities; such measures may, therefore, have some usefulness as outcomes in the evaluation of medical care for asthmatic patients. However, these measures may not adequately detect important differences in the number and duration of asthma attacks, the extent of wheezing between attacks, breathlessness, feelings of "tightness." Measurement of the latter phenomena may be necessary in order to precisely assess differences in health status for asthmatic patients. The optimal balance between the validity of health measures across a wide range of medical diagnoses and problems and the advantages of precise measurement of specific symptoms remains to be defined. The importance of a solution to this problem has been acknowledged and solutions are currently being investigated.⁴¹

The second validity issue concerns the extent to which interrelationships among health status measures are the same for subgroups within the general population. Since information about the relationships among health status measures is used to determine the validity of each measure, it is important to know the extent to which observed relationships are generalizable. Our preliminary findings regarding the generalizability of such validity findings across disadvantaged and nondisadvantaged groups suggest that important differences in the meaning of scores computed from health measures may exist.⁴² For example, measures of disability days are much less related to mental health concepts in disadvantaged groups than in nondisadvantaged groups. Thus, differences

in health as defined by disability days may reflect mental health phenomena to a greater extent for the nondisadvantaged than for the disadvantaged.

THE HIS HEALTH ASSESSMENT STRATEGY

One example of a health status assessment strategy in a large scale social experiment is that being used in the HIS. Selected groups of health variables being measured in the HIS are presented in Figure 1. Included are over 100 variables pertaining to physiologic and physical health, psychological health, social health and general health status perceptions. A variety of data gathering methods are being used, including (a) the observations of trained observers, (b) responses to questionnaires which are self-administered annually, and (c) biweekly summaries from health diaries.

Physiologic Health. Physiologic health pertains to the actual condition of the organ systems of the body. Observations regarding the actual condition of organ systems are being made by trained observers using 51 standardized procedures during multiphasic screening examinations. Included are laboratory tests of blood and urine samples, chest x-rays, and examinations of eyes, ears, teeth and skin.

Physical Health. Physical health pertains to the response of the organism (as a whole) to the condition of the organ systems. If the organ systems are deteriorating, physical health is the physical expression of that deterioration in the entire organism.

Functional Limitations due to health pertain primarily to problems with the physical functioning of respondents. Three kinds of activity which could be affected by limitations in physical functioning have been defined: 1) mobility, 2) physical activities, and 3) social role activities (including self-care activities). Social role includes activities typical for an individual of a given age. Physical activities pertain primarily to body movement capabilities such as ability to walk, stoop, or bend. Mobility refers to a person's ability to move around, within both the house and the community. As previously noted, one or

Figure 1

SELECTED HIS HEALTH MEASURES

Concept	Number of Scales	Type of Scale	Number of Items	Data Source
<i>Physiologic Health</i>	to be determined	-	51	multiphasic screening exam
<i>Physical Health</i>				
Functional Limitations				
Social Activity	1	Guttman	4	self-administered questionnaire
Physical Activity	1	Guttman	5	self-administered questionnaire
Mobility	1	Guttman	4	self-administered questionnaire
Physical Abilities	3	Likert	26	self-administered questionnaire
Disease Conditions (tracer method)	23	Weighted	313	self-administered questionnaire
Symptoms List	to be determined	-	27	self-administered questionnaire
Disability Days	-	-	-	bi-weekly modified diary
<i>Psychological Health</i>				
Anxiety	1	Likert	14	self-administered questionnaire
Depression	1	Likert	21	self-administered questionnaire
Behavioral/Emotional Control	1	Likert	3	self-administered questionnaire
<i>Social Health</i>	to be determined	Likert	11	self-administered questionnaire
<i>Health Perceptions</i>	5	Likert	32	self-administered questionnaire

more of these scale scores may reflect differences in psychological or social components of health in addition to physical functioning.

The Physical Abilities measures attempt to quantify the extent and severity of possible restrictions on the respondent's ability to engage in normal physical behaviors. Included are behaviors ranging from dressing oneself to engaging in sports such as swimming, bowling, and golf.

Disease Conditions are defined by over 300 items representing a subsample of 23 medical problems or diagnoses that can result in mild to severe disability. Excluded are conditions which rapidly lead to death. Included are those for which health care would be expected to be useful in terms of either alleviating symptoms or curing conditions. High blood pressure is an example of one condition which is included in this battery. Disease conditions are measured using the tracer method.⁴⁰

Disability Days reflect both sickness itself and the willingness of a person to adopt the sick role. The former concept is defined in terms of the extent to which a person feels able to function as well as usual. The latter concept is defined as the extent to which feelings of disability are translated into sick role behaviors (e.g., by reducing activities or by staying in bed).

The Symptoms List is directed at easily recognizable (generally acute) problems (such as colds, diarrhea or backache). These problems are expected to account for a large proportion of the morbidity a "healthy" person experiences over a period of time. Since most of the HIS population is expected to be "healthy," these common problems could potentially show the greatest differences among experimental groups.

Psychological Health. Psychological health is being assessed through a battery of items hypothesized to measure three concepts: 1) anxiety, 2) depression and 3) behavioral/emotional control. Items have been designed to allow construction of bipolar⁴³ scales for each concept as well as for general psychological well-being.²¹

Social Health. Social health pertains to several kinds of social participation variables, including: 1) nature and number of contacts with friends and relatives, 2) membership in voluntary groups, and 3) attendance at voluntary group meetings (including religious services). Respondents are also asked to rate how well they are getting along with others.

Health Perceptions. Health has also been defined as the respondent's own feeling or belief about personal health in general. For example, perception of personal health can be defined as the extent of agreement with statements like "I am not as healthy now as I used to be." The point of view of the respondent about his or her own health is unquestionably very important regardless of the extent to which perceptions are consistent with other information about health status.^{38,39} Thus, scales to measure health status perceptions are an important part of the HIS health assessment strategy. These scales measure perceptions regarding prior, current, and future health as well as current worry or concern about health and response to health state. The scores computed from these scales have a number of desirable properties: they are reliable, valid (in relation to a wide range of other health scores as well as use of health care services), and are among the least diagnosis-specific health measures.³⁸ Furthermore, because of the techniques that can be employed in constructing rating scales, their score distributions are not markedly skewed as is the case with most functional limitations and physical abilities measures. Thus, health perceptions scales are sensitive to differences in health for a much greater proportion of the population of interest.

Respondent Burden

Little is known about the effects of questionnaire length on data quality. The prevailing opinion is that the shorter the questionnaire the better. HIS questionnaires have been designed so that no more than

40 minutes will be required, on the average, to complete any one questionnaire booklet. Estimates of the actual average amount of time required do not exceed that standard. Available information about the amount of time required to complete individual sections or scales within a health questionnaire is most complete for the health status perceptions battery. Administration times for the health perceptions items have been recorded by trained interviewers in several studies; the 32 health status perceptions items require approximately seven and one half minutes to complete, on the average.^{38,39}

Studies of HIS Health Measures

Ongoing and proposed studies of the HIS health measures were designed to (a) maximize the reliability and validity of scores, (b) reduce, as much as possible, the number of variables necessary to define health, and (c) advance the state of the art of health measurement.

Estimates of the reliability of health scores are being obtained by studying the interrelationships among items within a scale and the stability of scale scores over time. The validity studies include analyses of the interrelationships among scale scores and correlations between scale scores and scores for the major health constructs, such as those derived by factor analysis. Attempts are being made to understand the extent to which these relationships are due to similarities and differences in both health components and methods of measurement. Relationships between health scale scores and scores for constructs other than health are also being examined using longitudinal and cross-sectional data gathered during the HIS. For example, the relationship between health scores and subsequent use of health care services is being examined in a study of the predictive validity of health scores.

A major goal of the HIS health measurement studies is an optimal reduction in the number of variables necessary to define health for purposes of hypothesis testing. It seems very doubtful that a single health index will suffice for most analytic purposes. The problem of

aggregating scores for various components of health has not been solved. Thus, it does not seem possible to translate a given combination of physical, psychological, and social health scores into one point on a single health continuum. Contrary to the notion of a single health index, the various components of health are not likely to be valued the same by everyone in a given population and a given component of health is not likely to be valued the same across subgroups of a given population.¹ It seems unlikely that the definition of health for a given person can be reduced to a single index score without tremendous loss of information.

SUMMARY

Several issues regarding the conceptualization and measurement of health status in policy relevant studies of medical care delivery have been discussed including: (a) reasons for health status assessment, (b) the nature and number of health concepts which should be measured, and (c) selected methodological considerations involved in developing a total measurement strategy.

Given that health status may not be related to either current notions regarding quality of care or use of health care services, direct measures of health status are needed in order to evaluate policy options regarding medical care delivery.

Health, which remains the most important human value for many people, should be treated as a multidimensional concept. Included are physical and physiologic, psychological, and social components.

The allocation of measurement resources among the various health components should correspond to the nature and extent of impact expected from present day optimal medical care. Current care impacts most on serious and nonserious morbidity. Thus, the largest amount of measurement resources should be allocated to measures of physical and physiologic health, followed by psychological and social health, respectively.

Measures of various health components are interrelated and in order to properly understand health outcomes, apparent health

outcomes should be interpreted with correlations among health components in mind. Furthermore, measures of the major health components should be constructed so as to define the positive states of health (well-being). Positive well-being involves not only the absence of negative states but also the attainment of positive states. Inclusion of positive states in the definition of health components marks the transition from a quantity to a quality of life measurement strategy.

Other considerations involved in the design of a total measurement strategy for health status assessment include: (a) the distributions of scores in the general population, (b) the reliability and validity of health scores in both disadvantaged and nondisadvantaged groups, and (c) respondent burden.

Scales to measure health status perceptions: (a) yield score distributions that are less skewed than distributions of scores for other scales, (b) are reliable and reasonably stable over time and (c) may be valid across a wide range of medical problems and diagnoses. The latter evidence suggests that measures of health status perceptions should be given greater emphasis in studies of medical care.

The Health Insurance Study (HIS), which The Rand Corporation is conducting for the Department of Health, Education and Welfare was offered as an example of a large scale social experiment in which a comprehensive approach to health status assessment is being employed. The HIS was designed primarily to estimate the effects of alternative cost-sharing arrangements on: (a) the use of health care services, (b) health status, and (c) quality of care. The omnibus health status assessment strategy designed for the HIS may have general applicability to problems involved in constructing and validating health measures for use in other studies.

REFERENCES AND NOTES

1. Ware, J. E. and Young, J. A. The Measurement of Health as a Value: Volume III of the Final Report, Contract No. HSM 110-72-299. Research Methods Branch, National Center for Health Services Research, DHEW, April, 1976.
2. Ware, J. E., Young, J. A., Snyder, M. K. and Wright, W. R. The Measurement of Health as a Value: Preliminary Findings Regarding Scale Reliability, Validity and Administration Procedures. Publication Number PB-239-616/AS, National Technical Information Service, Springfield, Virginia, 22151, 1974.
3. Using a modification of the Rokeach Value Survey (health was added) five independent studies of the importance of health in relation to other human values were conducted. Study participants (total N = 2050) were asked to rank health and 17 other human values "in order of their importance to YOU, as guiding principles in YOUR life." The median rank was computed for each of the 18 values and the 18 medians were ranked independently in each study. In every study, health was ranked highest (see references 1 and 2, above, for further information).
4. This comment is based on a personal communication from Dr. Robert H. Brook, who is at The Rand Corporation and the University of California at Los Angeles.
5. Kessner, D. M. and Kalk, C. E. A Strategy for Evaluating Health Services: Contrasts in Health Status, Volume 2. Washington D.C.: Institute of Medicine, National Academy of Sciences, 1973.
6. Kisch, A. I., Kovner, J. W., Harris, L. J. and Kline, G. A New Proxy Measure for Health Status, Health Services Research, Fall 1969, 223-230.
7. Suchman, E. A., Phillips, B. S., and Streib, G. F. An Analysis of the Validity of Health Questionnaires, Social Forces, 1958, 36, 223-232.
8. Friedsam, H. J. and Martin, H. W. A Comparison of Self and Physicians' Health Ratings in an Older Population, Journal of Health and Human Behavior, 1963, 4, 179-183.
9. Gurin, G., Veroff, J. and Feld, S. Americans View Their Mental Health, New York: Basic Books, 1960.

10. Srole, L., Langner, T. S., Michael, S. T., Opler, M. K. and Rennie, T. A. C. Mental Health in the Metropolis: The Midtown Manhattan Study, New York: McGraw-Hill, 1962.
11. Katz, M. M. and Lyerly, S. B. Methods for Measuring Adjustment and Social Behavior in the Community: I. Rationale, Description, Discriminative Validity and Scale Development, Psychological Reports, 1963, 13, 503-535.
12. Bradburn, N. M. The Structure of Psychological Well-Being. Chicago; Aldine Publishing Company, 1969.
13. Dohrenwend, B. P. Psychiatric Disorder in General Populations: Problem of the Untreated Case, American Journal of Public Health, June 1970, 1052-1064.
14. Goldberg, D. P., Cooper, B., Eastwood, M. R. et al. A Standardized Psychiatric Interview for Use in Community Surveys. British Journal of Preventive Social Medicine, 1970, 24, 18-23.
15. Dohrenwend, B. P. and Crandell, D. L. Psychiatric Symptoms in Community, Clinic, and Mental Hospital Groups. American Journal of Psychiatry, May 1970, 1611-1621.
16. Dohrenwend, B. P., Egri, G. and Mendelsohn, F. S. Psychiatric Disorder in General Populations: A Study of the Problem of Clinical Judgment. American Journal of Psychiatry, April 1971, 1304-1312.
17. Myers, J. K., Lindenthal, J. J. and Pepper, M. P. Life Events and Psychiatric Impairment. Journal of Nervous and Mental Disease, March 1971, 149-157.
18. Derogatis, L. R., Lipman, R. S., Covi. L., and Rickels, K. Neurotic Symptom Dimensions, Archives of General Psychiatry, May 1971, 454-464.
19. Berkman, P. L. Measurement of Mental Health in a General Population Survey. American Journal of Epidemiology, 1971, 94(2), 105-111.
20. Nussbaum, K. Objective Assessment of Degree of Psychiatric Impairment: Is it Possible? Hopkins Medical Journal, 1972, 133, 30-37.
21. Dupuy, H. J. Developmental Rationale, Substantive, Derivative and Conceptual Relevance of the General Well-Being Schedule, Draft Working Paper, June, 1973.
22. Schwartz, C. C., Myers, J. K. and Astrachan, B. M. Concordance of Multiple Assessments of the Outcome of Schizophrenia. Archives of General Psychiatry, October 1975, 1221-1227.

23. Tobin, S. S., and Neugarten, B. L. Life Satisfaction and Social Interaction in the Aging, Journal of Gerontology, 1961, 16, 344-346.
 24. Renne, K. S. Measurement of Social Health in a General Population Survey, Social Science Research, 1974, 3, 25-44.
 25. Greenblatt, H. N. Measurement of Social Well Being in a General Population Survey, Human Population Laboratory, California State Department of Health, 1975.
 26. Dohrenwend, B. S., Dohrenwend, B. P., and Cook, D. Ability and Disability in Role Functioning in Psychiatric Patients and Nonpatient Groups. In J. K. Wing and H. Hafner (Eds.) Roots of Evaluation. London: Oxford University Press, 1973.
 27. Fanshel, S. and Bush, J. W. A Health Status Index and Its Application to Health Services Outcomes, Operations Research, 1970, 18, 1021.
 28. Berdit, M. and Williamson, J. W. Function Limitation Scale for Measuring Health Outcomes, In R. L. Berg (Ed.) Health Status Indexes. Chicago: Hospital Research and Educational Trust, 1973.
 29. Patrick, D. L., Bush, J. W. and Chen, M. M. Toward an Operational Definition of Health. Journal of Health and Social Behavior, March 1973, 6-23.
 30. Reynolds, W. J., Rushing, W. A. and Miles, D. L. The Validation of a Function Status Index. Journal of Health and Social Behavior, 1974, 15(4), 271-288.
- Stewart, A. L., Ware, J. E. and Johnston, S. Construction of Scales Measuring Health and Health-Related Concepts from the Dayton Medical History Questionnaire, The Rand Corporation, Santa Monica, California, to be published.
32. Blischke, W. R., Bush, J. W. and Kaplan, R. M. Successive Intervals Analysis of Preference Measures in a Health Status Index. Health Services Research, 1975, 10, 181-198.
 33. Gilson, B. S. et al. The Sickness Impact Profile: Development of an Outcome Measure of Health Care. American Journal of Public Health, December, 1975, 1304-1310.
 34. Bergner, M., et al. The Sickness Impact Profile: Validation of a Health Status Measure. Medical Care, January 1976, 57-67.

35. Martin, D. K., Golson, B. S., Bergner, M. et al. Potential Use of a Health Status Assessment Instrument for Physician Training and Assessment of Outcome. Proceedings of the 14th Annual Conference on Research in Medical Education, Washington, D.C.: AAMC, November 1975, 277-282.
36. Studies of the interrelationships among ratings of the importance of 22 health concepts indicated the presence of three correlated health dimensions tentatively labeled Physical, Psychological and Social Health. These dimensions were observed in both disadvantaged and nondisadvantaged populations. Examples of items correlating highly with each factor (F) include (item-factor correlations shown in parentheses): Being free from pain (0.63 with F₁ labeled physical health, Having a healthy mind (0.72 with F₂ labeled psychological health, and Having good relationships with others (0.69 with F₃ labeled social health). See reference number one for more information.
37. Newhouse, J. P. A Design for a Health Insurance Experiment. Inquiry, March 1974, 5-27.
38. Ware, J. E. and Karmos, A. G. Development and Validation of Scales to Measure Perceived Health, Volume II of the Final Report, Contract Number HSM 110-72-299. Research Methods Branch, National Center for Health Services Research, DHEW, 1976.
39. Ware, J. E., Wright, W. R. and Snyder, M. K. Measures of Perceptions Regarding Health Status: Preliminary Findings as to Scale Reliability, Validity and Administration Procedures. Publication No. PB-242-726, National Technical Information Service, Springfield, Virginia, 22151.
40. Presence or absence refers to whether or not a person has a particular medical problem or diagnosis as evaluated, for example, vis-a-vis the tracer method (for an explanation of the tracer method, see reference number five, pp. 11-23). Severity refers to differences in health for persons who have a particular medical problem or diagnosis.
41. A study designed to develop measures which correlate with the final outcomes of care for given medical conditions, problems or diagnoses is currently being conducted by Doctors Robert H. Brook and Sheldon Greenfield, who are affiliated with The Rand Corporation and the University of California, Los Angeles.

42. Johnston, S. and Ware, J. E. Differences in the Meaning of Physical and Mental Health Scores Across Social Classes. The Rand Corporation, Santa Monica, California to be published.
43. A bipolar scale defines the extent of both positive and negative states along a single continuum. For example, the concept of "depression" is defined as the presence or absence of both "sadness" and "happiness."

