

Using Composite Health Status Measures to Assess the Nation's Health

PENNIFER ERICKSON, MS,* E. ALLEN KENDALL, MS,† JOHN P. ANDERSON, PHD,‡
AND ROBERT M. KAPLAN, PHD‡

Research in progress at the National Center for Health Statistics for evaluating the usefulness of composite measures of health status for assessing the nation's health is described. Three measures suitable for use in the general population, the Health Insurance Experiment-Functional Limitations (HIE-FL), the Health Utility Index (HUI), and the Quality of Well-being (QWB) scale, have been mapped to data collected in the 1980 National Health Interview Survey (NHIS). Analysis using current algorithms for making composite function status measures according to the QWB methods suggests that traditional single indicators of health tend to overestimate the level of health by about 10%. When symptoms and problems are added to the composite function score, the overestimate as measured by the single indicator is at least 50%. The authors are continuing to validate these algorithms, to develop similar ones for the HIE-FL and HUI, and to extend the analysis to data collected in 1977, 1979, and 1984. Current results indicate that to realize fully the benefits of composite measures, well-established, valid, and reliable measures of health-related quality of life should be included as part of the regular NHIS data collection procedures. Key words: health status assessment; health insurance; experiment-functional limitations; quality of well-being scale; health utility index; health-related quality of life. (Med Care 1989; 27:S66-S76)

At the national level, policy makers are faced with the need to balance the costs and benefits of various forms of health technology, for example, one drug versus another, one health program versus another, or a

drug versus surgery. In the past, health effects of interventions and treatments, especially at the national level, have been measured by condition-specific mortality and morbidity rates. More recently, as the pattern of illness has shifted to chronic diseases and people live longer but with more dysfunction, policy makers have made increasing use of indicators of limitation of activity and self-perceived health status.¹

These data are available through the National Center for Health Statistics (NCHS) and other sources, but they are of limited value in meeting the needs of policy makers when the benefit from interventions is quality of life associated with extended survival rather than that of quantity of life gained.

* From the Clearinghouse on Health Indexes, National Center for Health Statistics, Hyattsville, Maryland.

† From Social and Scientific Systems, Bethesda, Maryland.

‡ From the Division of Health Care Sciences, Department of Community Medicine, University of California, La Jolla, California.

Address correspondence to: Pennifer Erickson, MS, Chief, Clearinghouse on Health Indexes, Room 2-27, 3700 East-West Highway, Hyattsville, MD 20872.

One problem is that existing data do not give much information about the 85% of the population who report themselves to be unlimited in major activity or about the 50% who report themselves to be in either excellent or very good health.² As a result, the indicators that are presently collected give little insight about policies designed for "well" populations, defined as unlimited in major activity.

Another problem is that available indicators of health or well-being, such as self-perceived health status, are single measures. In the case of measures of a global concept, such as self-perceived health status, the indicator cannot be disaggregated.³ The significance of this for policy is that it is difficult to identify the cause and nature of the dysfunction and hence difficult to use the indicator to guide policy in an explicit way.

A third problem is that national data sets currently do not collect information about trade-offs—i.e., whether people would rather be in one health state or another.^{4,5} A major issue becomes that of survival in one state versus survival in another state, with both states having different definitions of function. As a result, traditional mortality and morbidity indicators, which reflect the prevalence of conditions but not the relative preference for living in given health states, provide limited information to policy makers faced with making these cost-benefit decisions. Thus, current indicators cannot give input to decision makers about individual's preferences for one program or treatment over another.

Recent research in assessing health status has addressed these limitations of traditional measures and their omission from existing data sets, especially in the context of health policy. Specifically, composite health status instruments have been developed that include several concepts of health-related quality of life. Among these are the following generic measures: Health Utility Index (HUI), Health Insurance Experiment (HIE) battery, Nottingham Health Profile,

Quality of Well-being (QWB) Scale, and Sickness Impact Profile (SIP).⁶⁻¹⁰ The Arthritis Impact Measurement Scale (AIMS), Health Assessment Questionnaire, and Kurtzke Disability Status Scale are examples of disease-specific measures.¹¹⁻¹³ The advantage of these composite measures, especially the generic instruments, is that they provide methods for obtaining more precise estimates of the health status of the 85% of the population with no limitation of activity.

Although these measures can provide a single score that assesses overall health status, they also include scores for subscales that represent different concepts of health-related quality of life. Thus, the total scores can be disaggregated so that the contribution of each subscale can be calculated and used for input into the policy decision.¹⁴

In addition, some composite measures include explicit estimates of the trade-offs that individuals are willing to make between different health states. Different methods have been used to collect this information so that it can be used at the population level. For example, Torrance used time trade-off, which is a variant of standard gamble, and expected utility theory.¹⁵ Others have used psychometric techniques such as magnitude estimation and category scaling.¹⁶⁻¹⁸

This study reports on a project to apply selected composite-measure methods to data collected in one of the general population surveys conducted by the NCHS, namely, the National Health Interview Survey (NHIS). The purposes of this retrospective data analysis are to evaluate the utility of these measures at the national level and to provide information on which to base recommendations for prospective data collection. This report describes procedures for mapping three composite measures to the NHIS and for compensating for missing data; it then presents national estimates of health status using one of the measures. The implications of using composite measures to arrive at estimates of the nation's health are discussed.

Methods

Selecting Composite Measures and NHIS Data

The NHIS is an on-going large scale health survey of the noninstitutionalized civilian population; the NHIS questionnaire consists of two parts, the core and special topic items. Questions in the core constitute approximately 70% of the questionnaire and are asked each year; these provide the basis for monitoring trends in population health levels over time. Among the core topics are limitation of activity, restriction in activity, prevalence of chronic conditions, and short-stay hospitalization.¹⁹ The remaining 30% of the questionnaire is devoted to special topics.

This format of a core plus special topic questions allows the NHIS to monitor trends in health status over time and to have flexibility in responding to issues of current and periodic interest. For example, health insurance coverage, cancer risk factors, and knowledge and attitudes about acquired immune deficiency syndrome (AIDS) were topics included in recent supplements. Of these, cancer and AIDS are new topics for NHIS, whereas health insurance coverage has been asked repeatedly since 1960.^{20,21}

The first step in selecting composite measures for use at the national level was to review the available research literature on the development and application of these measures in the context of the NHIS. We applied two criteria in this review. First, any measure selected for evaluation had to be appropriate for measuring health status in the general population. That is, the concept of health embodied in the measure had to be broad enough to fit the general population.²² Second, a candidate measure had to include items that were consistent with past NHIS data. This was important because in a retrospective analysis such as this, the composite measure could be constructed only if past data existed. Also, if the measure is to be included in future surveys, the data

should be consistent with those collected previously so that trends reported by NHIS can be preserved.

Based on these criteria, we selected three measures for evaluation: the HIE Functional Limitations questionnaire, the HUI, and the QWB. The HIE Functional Limitations measure (HIE-FL) was part of a larger health status battery developed for a large-scale social experiment to determine the health effects of a national health insurance program.^{7,23-26} Although many different concepts of health status and health-related quality of life were assessed as part of a strategy for determining if free care improved health status, only the functional limitations questionnaire had sufficient commonalities with the NHIS to be considered as part of this evaluation. The HIE-FL assesses four concepts of health status or health-related quality of life: 1) mobility, 2) role functioning, 3) physical functioning, and 4) self-care.

The HUI is a scheme for classifying individuals into mutually exclusive health states.²⁷ This measure was developed for resource allocation purposes and has been used to assess the effectiveness of neonatal intensive care units, tuberculosis screening programs, and renal dialysis.^{28,29} Like the HIE-FL, the HUI includes four concepts of health-related quality of life: 1) physical functioning, mobility, and physical activity; 2) role functioning, including self care and role activity; 3) social-emotional function, including emotional well-being and social activity; and 4) health problems.

The QWB Scale is also a classification scheme, but it uses a standardized questionnaire to collect data by which individuals can be assigned to a health state.³⁰ Since its development, the QWB has been used with arthritis patients and persons with chronic obstructive pulmonary disease as well as for the evaluation of health promotion programs for the elderly.³¹⁻³³ The QWB collects data on four concepts of health-related quality of life: 1) physical ac-

TABLE 1. Categories of Agreement Between NHIS and Research Instruments: Three Examples

Definite agreement
HUI: being able to get around the house, yard, neighborhood, or community without help from another person
NHIS: walking, going outside, no reported problems
Reasonable agreement
HIE-FL: in bed or chair most or all of the day
NHIS: because of health or disability does anyone stay in bed most or all of the day
No agreement
QWB: most or all of day in chair or couch
NHIS: no correspondence

HUI: Health Utility Index, NHIS: National Health Interview Survey, QWB: Quality of Well-being Scale.

tivity, 2) mobility, 3) social activity, and 4) symptoms and problems. The model and method for collecting data for the QWB were initially developed to provide insights for health planning decisions in New York State.³⁴

After identifying the three measurement strategies that met the criteria, we reevaluated recent NHIS questionnaires for similarities to the research approaches. We selected four NHIS data years, 1977, 1979, 1980, and 1984, for retrospective analysis because the data collected in these years were most comparable to the three research measures; for 1979 and 1980, the same NHIS questionnaire was used. In addition to providing the opportunity for calculating composite scores for each year, these multiple data points allow for a trend analysis over almost a 10-year period.

Evaluating Comparability

We examined each of the four NHIS questionnaires, including both core and special topic questions, for items that agreed with or matched those in each of the research instruments. This consisted of identifying specific items asked in the NHIS that seemed to be assessing the same concepts of health-related quality of life that were in-

cluded in each of the research instruments. This matching procedure generated four comparability tables, one for each of the concepts of health-related quality of life included in each of the research instruments. Within each of these tables, items in the NHIS questionnaires were grouped into one of three categories: 1) definite agreement, 2) reasonable agreement, and 3) no agreement.

Comparisons of sample items from the research instruments with items from the 1979 and 1980 NHIS are given in Table 1 to illustrate the category definitions; these examples have been drawn from the concepts in the research instruments that relate to physical activity or physical function. As indicated by the comparison of items from the HUI and the NHIS, a categorization of "definite agreement" was made even if the wording was not identical. Instead, we looked for questions that seemed to assess the same underlying health concept.

A categorization of "reasonable agreement" was made for questions that seemed to measure related concepts of health but lacked a one-to-one correspondence between the concepts in each item. In Table 1, the example is between items in the HIE-FL and the NHIS; both items inquire about a respondent's being in bed most or all of the day; however, the 1979 and 1980 NHIS collected no information on persons who were chairbound.

For some items in each research instrument the NHIS had no corresponding information. The Table 1 example of "no agreement" is for the QWB and the NHIS. As was noted above, the NHIS does not collect information on being either in a chair or couch for most of the day.

Each question was also examined for recall period, for mode (i.e., actual performance of an activity or reported capacity to perform an activity), and for selected respondent characteristics (e.g., age). The research instruments tended to be consistent throughout with regard to both recall and mode. For example, the QWB collects infor-

mation on a day-by-day basis over a 4-day recall period using the performance mode. On the other hand, questions within any given NHIS data year varied with regard to both recall and mode from one section of the questionnaire to another. For most variables relevant to this analysis, including restricted activity days, the NHIS asks about a 2-week recall period. For the subscales, restricted activity days were used to estimate the duration of time that a respondent spent in each relevant health state. However, for some functional limitations NHIS did not specify the recall period, as for example the items on activities of daily living (ADL) that were included in the 1979 and 1980 special topic section on home care. Because NHIS defines these limitations as chronic, we assumed that they existed for 14 days.

Differences in respondent characteristics were also noted on the comparability tables. For example, as part of its subscale on mobility, the QWB asks about driving a car for persons aged 16 years and older; the 1979-1980 NHIS asked for similar information but used age 17 as the lower age limit. Although most of the deviations in respondent characteristics came from the use of different age cutoffs between the research instrument and the NHIS, different approaches to health survey design also caused some discrepancies. For example, in 1979 and 1980 the NHIS asked about ability to drive or use public transportation only of persons who indicated that they were dependent in performing ADLs or in "instrumental" ADLs.

Imputation

For items with either no or reasonable agreement, we developed various strategies for imputing the missing information. The most extensive of these strategies is that designed to estimate the occurrence of symptoms in the NHIS sample. For the majority of the symptoms and health problems used in both the HUI and the QWB, the NHIS has

no corresponding information; instead of collecting information on symptoms, the NHIS asks about more than 100 self-reported health conditions using six different lists.

For this study, we used the NHIS information on conditions to develop estimates for the symptoms and health problems used in the HUI and QWB. Specifically, data from the 1980-81 National Ambulatory Medical Care Survey (NAMCS) provided a mapping between self-reported health conditions and symptoms. The NAMCS survey instrument includes a physician's diagnosis of a patient, i.e., the health condition, and the reason given by the patient for the visit to the physician's office, i.e., the symptom.

In the imputation for the QWB, we recoded NAMCS data to conform, on the one hand, to the QWB symptom-problem classification (Table 2) and, on the other hand, to the NHIS condition codes. We then prepared tables showing for each NHIS condition the proportion of individuals who fell into each symptom-condition group by age and sex. Because differences by age and sex were not statistically significant, we used data from the total population to form probability distributions of symptoms associated with each condition. A potential bias was created because the NAMCS is a sample of individuals who sought medical care, rather than being a probability sample of all individuals in the U.S. population. To identify and correct for this bias, a practicing clinician reviewed the symptom distributions for validity; when necessary the distributions were adjusted to account for differences between the NAMCS and NHIS populations.

The NAMCS distribution for the common cold is shown in Table 3. This shows that an individual reporting a cold has a 62% chance that symptom-problem category 18 will be assigned. When an NHIS respondent reported more than one condition, we followed the standard QWB practice of assigning the "most undesirable" symptom or

TABLE 2. Symptoms and Health Problems Used in the Quality of Well-being Scale

Number	Description
1	Death (not on respondent's card)
2	Loss of consciousness such as seizures, fainting, or coma (out cold or knocked out)
3	Burn over large areas of face, body, arms, or legs
4	Pain, bleeding, itching, or discharge from sexual organ (does not include normal menstrual bleeding)
5	Trouble learning, remembering, or thinking clearly
6	Any combination of one or more hands, feet, arms or legs either missing, deformed, paralyzed, or broken (includes wearing artificial limbs or braces)
7	Pain, stiffness, weakness, numbness, or other discomfort in chest, stomach, side, neck, back hips, or any joints of hands, feet, arm or legs
8	Pain, burning, bleeding, itching, or other difficulty with rectum, bowel movements, or urination
9	Sick or upset stomach, vomiting or loose bowel movements, with or without fever, chills, or aching all over
10	General tiredness, weakness, or weight loss
11	Cough, wheezing, or shortness of breath with or without fever, chills, or aching all over
12	Spells of feeling upset, being depressed, or of crying
13	Headache, or dizziness, or ringing in ears, or spells of feeling hot, or nervous, or shaky
14	Burning or itching rash on large areas of face, body, arms, or legs
15	Trouble talking, such as a lisp, stuttering, hoarseness, or being unable to speak
16	Pain or discomfort in one or both eyes or any trouble seeing after correction
17	Overweight for age and height, or skin defect of face, body, arms, or legs, such as scars, pimples, warts, bruises, or changes in color
18	Pain in ear, tooth, jaw, throat, lips, tongue; several missing or crooked permanent teeth—includes wearing bridges, or false teeth; stuffy, runny nose; or any trouble hearing—includes wearing a hearing aid
19	Taking medication or staying on a prescribed diet for health reasons
20	Wore eyeglasses or contact lenses
21	Breathing smog or unpleasant air

problem from the set of those imputed for each of the conditions.

In the NHIS, each condition list is given to a one-sixth subsample of the total sample; thus, conditions reported by each respondent are weighted by a "correction factor" of approximately 6. This design allows the NHIS to derive national estimates of the prevalence of chronic conditions occurring in either the total or selected subgroups of the population in any given year. However, this sampling scheme limits the analyses that can be done using the individual as the unit of analysis, since information is unavailable about the presence of certain comorbidities. For example, for persons who were asked about conditions affecting hearing, speech, or vision (which appeared on List 4 in the 1980 NHIS), no estimate is available as to whether they also had arthritis or other conditions affecting the bone and muscle (List 2 in the 1980 NHIS). One implication of this is that the indication of the most undesirable symptom is underrepresented, which yields a biased composite score for the QWB. We have not yet developed an approach to compensate for this bias.

The final part of the symptom and problem imputation for the QWB was to assign a duration; not all health conditions cause symptoms all of the time. Different approaches to imputing these data were examined. The current approach links the duration of symptoms and problems to the number of days of limitations on the QWB Social Activity Scale (SAC). This scale reports limitation in activity during the 2-week period used; the NHIS variables used to form the social activity subscale were similar to the QWB variables for this subscale. As a first attempt at imputing duration, we assumed that symptoms would be present for 2 days more than the duration of the limitation in social activity. Thus, duration of any symptom or problem was imputed to range from 2 to 14 days. We are continuing to refine this

arbitrary approximation to obtaining duration of symptoms.

Results

To date, our efforts have focused on developing procedures for creating composite measures according to the methodology used by the QWB, which consists of three functional subscales (mobility, physical activity, and social activity) plus a list of symptoms and health problems. Because the algorithms that we have developed for forming the subscale and overall scores are still being validated, we are presenting percentages of persons without dysfunction rather than calculated scores.

For the total U.S. population, the percentage of persons within a given age and sex group with no dysfunction on each of the functional subscales is about the same; for example, for males less than 17 years old, the percentages with no dysfunction in mobility, physical activity, or social activity were 87, 86 and 87, respectively. The percentage with no dysfunction decreases with age for both males and females; approximately 50% of persons 65 years and older have some dysfunction in mobility, physical activity, or social activity.

Of the persons with no limitation in social activity who are less than 65 years old, 98% also have no limitations in the two other function subscales. For persons 65 years and older, 96% of the males and 97% of the females with no limitation in social activity also have no limitation in mobility and physical activity.

Figure 1 shows a comparison of the percentages of persons in the total U.S. population with no dysfunction as measured by the following: 1) the traditional NHIS limitation of activity indicator; 2) a composite function status indicator that combines mobility, physical activity, and social activity; and 3) the overall QWB score. The highest percentage of persons with no dysfunction is obtained by using the single indicator, limitation of activity; percentages ranged

TABLE 3. Percentage of Symptoms and Problems Associated with the Common Cold Based on Data from the National Ambulatory Care Survey

Number	Symptoms and Problems	%
7	Pain, stiffness, numbness, or other discomfort in torso, neck or joints	1
11	Cough, wheezing, or shortness of breath	15
13	Headache, dizziness, or spells of feeling hot, or nervous or shaky	4
16	Pain or discomfort in one or both eyes (such as burning or itching) or any trouble seeing after correction	1
18	Pain in ear, throat, or stuffy, runny nose, or trouble hearing standard (average) symptom or problem	62
		17

from 96% for persons less than 17 years old to 55% for persons 65 years and older.

Percentages of persons with no dysfunction are lower for the composite function status indicator, ranging from 86% for persons less than 17 years old to 51% for those 65 years and older. The overall QWB score produced the lowest percentage of persons with no dysfunction. When symptoms and problems are added to the composite function measure, the percentage with no dysfunction ranges from 65% to 2% for persons less than 17 years old to those 65 years and older, respectively. The trend for all three measures was the same for males and females, although a slightly higher percentage of females have dysfunction than do males.

Similar data are shown for the more than 26 million persons in the U.S. who reported having arthritis in 1980 (Fig. 2). Because fewer than 1% of the persons with arthritis are less than 17 years old, we excluded this age group from this analysis. Also, because of the methods used to impute symptoms as discussed above, we assigned all of the persons with arthritis some symptom for at least 2 days. As a result, everyone with this condition had some dysfunction; put another way, for each age group, 0% of per-

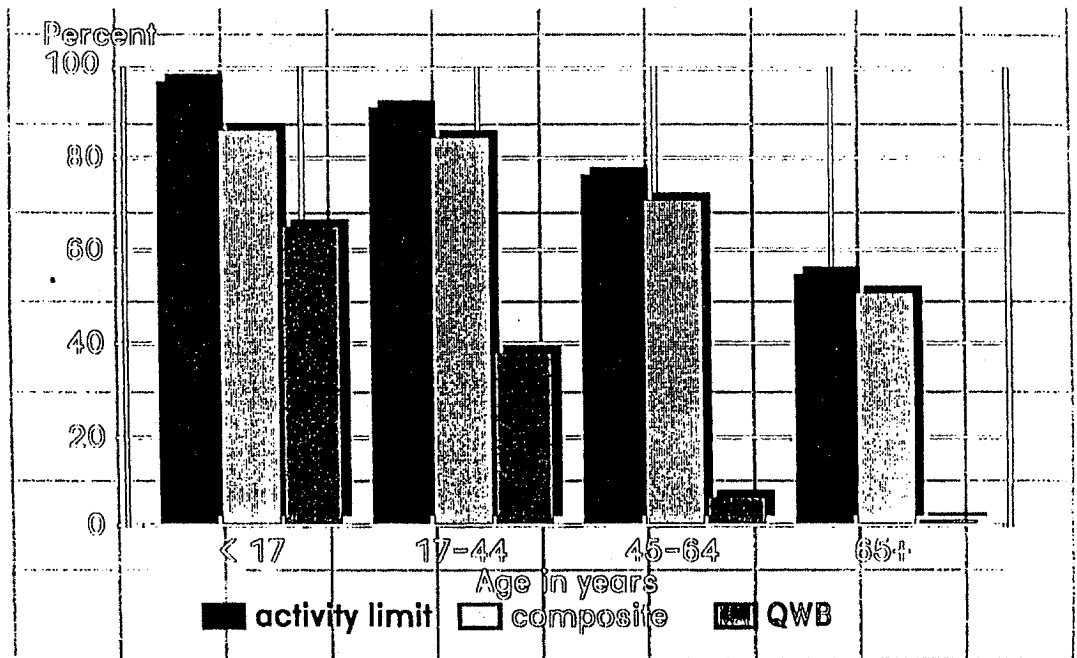


FIG. 1. Percent of persons with no dysfunction. U.S. population, NHIS 1980.

sons had no dysfunction as indicated by the QWB.

Using the traditional activity limitation indicator, the percentage of persons with arthritis having no dysfunction ranged from 71% to 44% for persons 17 to 44 years old to 65 years and older, respectively. Although the percentages of persons with no dysfunction are lower for the composite function status measure, the pattern across age groups is similar to that for limitation of activity. As with the total population (i.e., regardless of health condition), the distributions for males and females were similar, although females reported a slightly higher dysfunction than did males.

For the total U.S. population, the activity limitation indicator gives an approximately 10% higher estimate of the nation's health than does the composite function status measure; the overestimate is 12% for persons younger than 17 years and 8% for persons 65 years and older. For persons with arthritis, the overestimate arising from the

use of a single indicator rather than a composite score is greater for younger than for older persons. Limitation of activity overestimates health status by 22% for persons 17 to 44 years old and by 10% for those 65 years and older. When symptoms are included, i.e., when the overall QWB score is used, limitation of activity presents an even larger overestimate of health status; for the total U.S. population regardless of age, the overestimate is at least 50%.

Discussion

Our results suggest that computing composite measures for the nation's health is possible from data previously collected in the NHIS. However, measures developed from retrospective analyses are necessarily based on several assumptions and, in our case, imputations for incomplete data. We are currently assessing the validity of the algorithms, including the imputations, that we have developed. This includes compar-

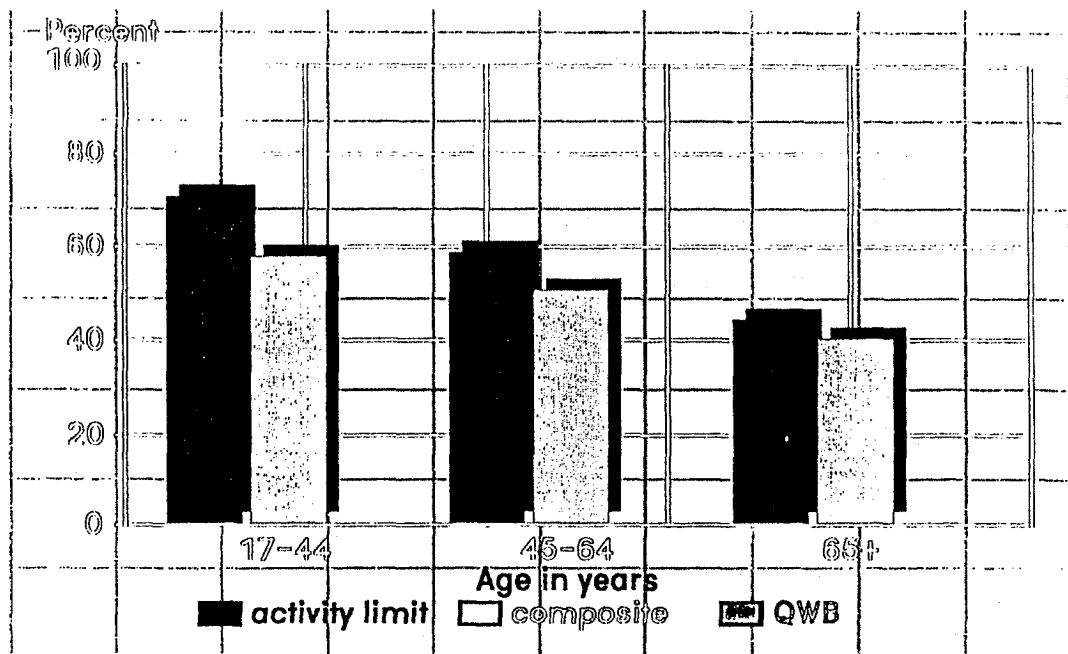


FIG. 2. Percent of persons with no dysfunction. Persons with arthritis, NHIS 1980.

ing scores based on NHIS data with those obtained in other populations in which the research instruments have been used in their original form. For example, comparison of scores obtained for the original QWB and for the NHIS-QWB among persons with arthritis will test the face validity of our algorithms for this research instrument. After thorough validity testing, investigators should be able to use scores rather than percentages (as done in this report); then, the NHIS-QWB can be used in the same way as the original QWB has been.

These findings, based on preliminary algorithms, indicate that single measures tend to overestimate the level of the nation's health. Although the magnitude of the overestimate varies by age group and choice of composite measure, it is at least 7% for the total population. For arthritis, selected as an example of one of the more prevalent chronic conditions reported in the NHIS, the overestimate is at least 10%. Including symptoms and health problems in the com-

posite score greatly increases the overestimate of health status as given by the single activity limitation indicator. This suggests that symptoms and problems are very prevalent in the population and that health indicators that exclude them will upwardly bias estimates of population health status.

Although our results indicate that composite health status measures can be calculated from existing NHIS data, several problems exist. One, which arises from limitations in NHIS design, is the absence of data on symptoms and problems. Owing to imputations of symptom-problem complexes from NAMCS data, the relationship between symptoms and other variables is statistical rather than actual. As a result, the NHIS-QWB can be calculated only for population subgroups, not for individuals. This limits the types of policy analyses that can be performed to those with sufficient population subgroup sizes; sizes needed for analyses are larger than they would be if the data had been obtained at the individual

level. A closely related problem is the one-sixth subsample design; because of this, not all symptom data are available for all persons in the sample. This results in an overestimate of health status using the constructed NHIS-QWB.

Another problem is the lack of available preference weights to quantify tradeoffs between functional subscales and symptom-problem complexes. Currently, data from a general population survey conducted in San Diego, California, in 1974 and 1975 are being used to combine the four concepts of health-related quality of life into an overall QWB score. One empirical question is whether the preference weights from the California survey represent those for the total U.S. population. Another is whether these weights, assuming that they are representative, have changed over time. More empirical research is needed on the collection and interpretation of preferences.

Despite these problems, the benefits of using composite measures of health-related quality of life to assess national health levels should not be underestimated. As discussed above, one major advantage is a more comprehensive estimate of the nation's health status. Another is that the NHIS data can be put to wider use. For example, combining composite scores such as those from the QWB with life expectancy to calculate an estimate of quality-adjusted life years (QALY) allows for comparison across disease and other problem-specific target populations. Thus, they can be used in resource allocation models and evaluation studies, uses that go beyond the current applications of NHIS data. In addition, QALY data improve interpretation of trends as indicated by the analyses by Verbrugge and Manton and Soldo.^{35,36} The usefulness of composite measures as less biased estimates of population health levels and as components in estimates of QALYs argues for including well established, valid, and reliable measures of health-related quality of life as part of the regular NHIS data collection procedures.

Acknowledgments

Lester Curtin, Thomas F. Drury, Theodore Ganiats, Susan Jack, Abigail Moss, Andrea Rauh, Ildy Shannon, George Torrance, and John E. Ware, Jr. all provided valuable assistance at various points in the development of this project. Any misinterpretations or errors remain the responsibility of the authors.

References

1. NCHS. Health, United States, 1984. DHHS Publication No. (PHS) 85-1232. Hyattsville, Maryland: National Center for Health Statistics, 1984.
2. Patrick DL, Erickson P. Assessing health-related quality of life in general population surveys. *Vital Health Stat* [2]. Hyattsville, MD: National Center for Health Statistics (in forthcoming).
3. Ware JE. Methodological considerations in the selection of health status assessment procedures. In: Wenger NK, et al. *Assessment of quality of life in clinical trials of cardiovascular therapies*. New York: LeJacq Publishing 87.
4. Lipscomb J. Value preferences for health: meaning, measurement, and use in program evaluation. In: Kane RL, Kane RA, eds. *Values and long-term care*. Lexington, MA: Lexington Books, 1982:27.
5. Torrance GW. Multiattribute utility theory as a method of measuring social preferences for health states in long-term care. In: Kane RL, Kane RA, eds. *Values and long-term care*. Lexington, MA: Lexington Books, 1982:127.
6. Torrance GW, Thomas DL, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Serv Res* 1972;7:118.
7. Brook RH, Ware JE, Davies-Avery A, et al. Overview of adult health status measures fielded in RAND's Health Insurance Study. *Med Care* 1979;17(suppl 1).
8. Hunt SJ, McEwen J. The development of a subjective health indicator. *Sociol Health Illness* 1980;2:231.
9. Kaplan RM, Bush JW. Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychol* 1982;1:61.
10. Bergner M, Bobbitt RA, Carter WB, et al. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981;19:787.
11. Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis: the Arthritis Impact Measurement Scales. *Arthritis Rheum* 1980;23:146.
12. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. *J Rheumatol* 1982;9:789.
13. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale. *Neurology* 1983;33:1444.
14. Moriyama IW. Problems in the measurement of health status. In: Sheldon EB, Moore WE, eds. *Indicators of social change*. New York: Russell Sage Foundation, 1968:573.

15. Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socioecon Plann Sci* 1976;10:129.
16. Carter WB, Bobbitt RA, Bergner M, et al. Validation of an interval scaling: the Sickness Impact Profile. *Health Serv Res* 1976;11:516.
17. Hunt SM, McKenna SP, McEwen J, et al. A quantitative approach to perceived health status: a validity study. *J Epidemiol Commun Health* 1980;34:281.
18. Patrick DL, Bush JW, Chen MM. Toward an operational definition of health. *J Health Soc Behav* 1973;14:6.
19. NCHS. Current estimates from the National Health Interview Survey, United States, 1986. DHHS publication no. (PHS) 87-1592. (*Vital Health Stat* [10] Hyattsville, MD: National Center for Health Statistics, 1987.
20. Dawson DA, Thornberry O. AIDS knowledge and attitudes for November: provisional data from the National Health Interview Survey. *Advancedata* 1988;150:1.
21. NCHS. The National Health Interview Survey design, 1973-84, and procedures, 1975-83. DHHS Publication No. (PHS) 85-1320. (*Vital Health Stat* [1] Hyattsville, MD: National Center for Health Statistics, 1985.
22. Bergner M, Rothman ML. Health status measures: an overview and guide for selection. *Annu Rev Public Health* 1987;8:191.
23. Brook RH, Ware JE, Rogers WH, et al. Does free care improve adults' health? results from a randomized controlled trial. *N Engl J Med* 1983;309:1426.
24. Newhouse JP. A design for a health insurance experiment. *Inquiry* 1974;11:5.
25. Newhouse JP, Manning W, Morris C, et al. Some interim results from a controlled clinical trial of cost sharing in health insurance. *N Engl J Med* 1981;305:1501.
26. Ware JE, Brook RH, Davies-Avery A, et al. Conceptualization and measurement of health for adults in the Health Insurance Study, vol I. Model of health and methodology. R-1987/1-HEW. Santa Monica, CA: The RAND Corporation, 1980.
27. Drummond MF, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford, England: Oxford University Press, 1977.
28. Boyle MH, Torrance GW, Sinclair JC, et al. Economic evaluation of neonatal intensive care of very-low-birth-weight infants. *N Engl J Med* 1983;308:1330.
29. Torrance GW, Sackett DL, Thomas WH. Utility maximization model for program evaluation: a demonstration application. In: Berg RL, ed. *Health status indexes*. Chicago, IL: Hospital Research and Educational Trust, 1973:156.
30. Bush JW. General health policy model/Quality of Well-being (QWB) scale. In: Wenger NK, et al, eds. *Assessment of quality of life in clinical trials of cardiovascular therapies*. New York: Lejacq Publishing, 1984:189.
31. Balaban DJ, Sagi PC, Goldfarb NI, et al. Weights for scoring the Quality of Well-being instrument among rheumatoid arthritics: a comparison to general population weights. *Med Care* 1986;24:973.
32. Kaplan RM, Atkins CJ, Timms R. Validity of a Quality of Well-being scale as an outcome measure in chronic obstructive pulmonary disease. *J Chronic Dis* 1984;37:85.
33. Kincade JE, Patrick DL, Beresford SA, et al. Functional status and perceived quality of life among older persons. In: *Proceedings of the 1987 Public Health Conference on Records and Statistics* DHHS Publication No. (PHS) 88-1214. Hyattsville, MD: National Center for Health Statistics, 1987.
34. Bush JW, Chen MM, Patrick DL. Health status index in cost effectiveness: analysis of PKU program. In: Berg RL, ed. *Health status indexes*. Chicago, IL: Hospital Education and Research Trust, 1973:172.
35. Verbrugge IM. Longer life but worsening health? trends in health and mortality of middle-aged and older persons. *Milbank Mem Fund Q* 1984;62:475.
36. Manton KG, Soldo BJ. Dynamics of health changes in the oldest old: New perspectives and evidence. *Milbank Mem Fund Q* 1985;63:206.