The SF-12 as a Population Health Measure: An Exploratory Examination of Potential for Application

James N. Burdine, Michael R. J. Felix, Amy Llewellyn Abel, Charles J. Wiltraut, and Yvette J. Musselman

Objective. To describe the relationships among functional health status measures (SF-12 physical and mental components summary scores), traditional measures of community health status, and social determinants of health among respondents to community health status surveys conducted in nine different communities.

Data Sources/Study Setting. Data collected as part of comprehensive community health status assessments conducted in each of nine communities (in seven states) between 1992 and 1997. The purpose of each assessment was to gather data to plan and evaluate population health improvement initiatives.

Study Design and Data Collection. This is an opportunistic study drawing on the universe of community health survey data collected by the authors to support local health improvement initiatives. Both community-level as well as an aggregate of individual-level measures are used in the analysis. Within each locality, survey respondents were randomly selected using a telephone-facilitated, mailed survey methodology.

Principal Findings. The key variables reported here are functional health status measures (SF-12) and social determinants of health variables. SF-12 physical and mental component scales correlated with two of four traditional measures of community health status. At the aggregate level of analysis, significant relationships were found for seven of nine social determinants of health measures when compared with SF-12 component summary scores. Relationships between social determinants measures and PCS-12 and MCS-12 scores suggest both application possibilities and the need for additional analysis in order to understand the nature of those relationships.

Conclusions. Physical and mental health functioning summary scores as measured by the SF-12 are useful in describing overall community health status when compared with traditional measures such as total deaths, age-adjusted mortality, or physician to population ratio. The SF-12 can also be used to measure the relationship between physical and mental health functioning (as proxies for community health status) and the social determinants of health. This analysis can help to refine our understanding of how social determinants and health status interact in a community or population as a precursor to the development of models of community or population health.

Key Words. SF-12, outcomes, health status assessment, community, population health

Measurement of the outcomes of clinical care has been developing as both a field and a science over the past two decades. Beginning with the RAND Health Insurance Experiment, through the Medical Outcomes Study, the desire and ability to quantify medical outcomes has increased dramatically. The standardized measurement of health outcomes, through instruments such as the SF-36, and more recently the SF-12, has had significant benefit for all fields and professions concerned with health (Tarlov, Ware, Greenfield, et al. 1989; Felix and Burdine 1995; Ware, Kosinski, and Keller 1996b; Joint Commission for the Accreditation of Healthcare Organizations 1997; Institute of Medicine 1998).

A body of work in community and population health improvement has focused on translating the tools and approaches for measuring medical outcomes in clinical contexts to measuring health outcomes in community contexts, specifically through the use of instruments such as the SF-36 and SF-12 (Eden 1998; McHorney, Kosinski, and Ware 1994; Tarlov, Ware, Greenfield, et al. 1989; Ware, Kosinski, and Keller 1996a; Ware, Bayliss, Rogers, et al. 1996; Medical Outcomes Trust 1995). The desired result of this application is the ability to measure population health status, its correlates, and their changes over time in order to plan, implement, and evaluate population or community health status interventions. Use of the SF-12 for measuring health status in this process links the advances that have been made in the standardization and measurement of health status in enrolled or patient populations using this tool, to the possibility of application for community health status measurement. Creating this linkage could represent substantial gains in community health measurement because of the non-medical factors that play an important role in determining the health status of individuals and populations; these factors have been called the social or broader determinants of health (Kaplan and Lynch 1997; Steckler, Allegrante, Altman, et al. 1995; Amick et al. 1995; Gold and Stevens 1998; Illich 1976; LaLonde 1975; Blum 1974). Effecting changes in these social determinants is key to the improvement of population health, and to effect this change the relationships

---

These data were collected under contracts between the authors' organization and community partnerships or institutions in each individual community.

Address correspondence to James N. Burdine, Dr.P.H., President and Senior Scientist, Felix, Burdine and Associates, 5100 Tilghman Street, Suite 215, Allentown PA 18104. Michael R. J. Felix is Chief Executive Officer; Amy Llewellyn Abel, M.S.P.H. is Vice President; Charles J. Wiltraut is Associate; and Yvette J. Musselman is Research Coordinator, all with Felix, Burdine and Associates. This article, submitted to *Health Services Research* on August 31, 1998, was revised and accepted for publication on September 15, 1999.

among and between the social determinants and measures of population or community health must be defined. There has been a history of population-based research on the effects of select social determinants, such as income, employment, and social class on health status, as measured by indicators such as mortality (Marmot, Bobak, and Smith 1995; Wilkinson 1986; Evans 1994). In this article, we will use functional status as the indicator to describe the relationships between some of these social determinants and population health status.

First, we describe a method for measuring population health through a series of community-based health status assessments conducted across the United States between 1992 and 1997. Each assessment incorporates the SF-12, an instrument typically used to measure the health progress or decline of individuals in a clinical setting. We then validate the use of the SF-12 as a measure of community health status by describing the relationship between this and other traditional measures of community health status. Third, using aggregate data, we explore the relationships between functional health status and variables that capture the social determinants of health. Based on the results obtained from analysis, we propose an application for the SF-12 as a measurement tool for population health status and raise questions for future research.

METHODS

Between 1992 and 1997, we assisted numerous local partnerships and coalitions in communities across the country with conducting population health status assessments. One component of each of those assessments was a general population survey. The data from nine of these community surveys are used in the analysis reported here.

Data Collection

The standard recruitment process for each survey included facilitating telephone calls placed based on random-digit-dialing lists obtained from commercial list vendors. Professional telephone interviewers asked to speak to the adult with the next birthday to screen potential respondents for age (18 or older) and to compensate for gender-biased telephone response behavior. If that person was unavailable, the interviewer scheduled a callback to that person.

When reached, potential respondents were asked if they would be willing to participate in a community health study by completing a survey

that would be mailed to them within one week. Each respondent was offered a two dollar cash incentive during the phone call to participate in the study. Calls were made between 4:00 and 9:00 p.m. on weekdays and 1:00 to 9:00 p.m. on weekends, local time.

Respondents who agreed to participate in the survey were mailed a packet with a personalized letter thanking them for their cooperation, a two dollar bill, a survey booklet, and a postage-paid return envelope. This approach was developed and refined on the basis of our experience and reflects the findings of others collecting functional health status data (William 1990; McHorney, Kosinsky, and Ware 1994).

Survey Instrument

The survey booklets for each survey were 26 to 28 pages, required approximately 30 to 40 minutes to complete in pretesting, and assumed a seventh grade literacy level. In communities with a significant Spanish-speaking population (>5 percent), Spanish translations of the survey were distributed based on the language spoken by the respondent during the recruiting telephone call. Local modifications in translation were made to accommodate the community (e.g., Mexican Spanish versus Puerto Rican Spanish).

The instrument used in the community survey process has been continuously refined since its development in 1991. The SF-12 functional health status instrument is a component of each of the community surveys. Other data elements in the survey instrument include questions about acute and chronic conditions; behavioral risk factors; receipt of preventive screenings; perceptions of access to primary medical care as well as its quality; perceptions of community issues; utilization of medical, health, and human services; insurance coverage; and other demographic characteristics. Questions in the survey have their origins in the Medical Outcomes Study, the Centers for Disease Control and Prevention Behavioral Risk Factor Survey (CDC-BRFSS), the Primary Care Assessment Survey (PCAS), and the Group Health Association of America's Consumer Satisfaction Survey. All other elements were developed by us and in conjunction with the communities that have participated in the surveys since 1992.

As shown in Table 1, telephone cooperation rates varied from 52 percent to 76 percent, averaging 60 percent across the nine communities. Response rates to the mailed survey component ranged from 58 percent to 82 percent, averaging 71 percent across the nine surveys. The overall response rate for each survey ranged from 37 percent to 62 percent, with an average total response rate of 47 percent (calculated by multiplying the telephone

Table 1: Survey Site Characteristics

Site	*Type**	*Total Population*	*Population Density (per Square Mile)*	*Year of Survey*	*Sample Size (n)*	*Phone Cooperation Rate (%)*	*Mail Response Rate (%)*	*Overall Response Rate (%)*
1	r, s, u	540,000	761	1992	2,446	75	82	62
2	r	87,000	86	1994	2,111	62	82	51
3	r	41,000	37	1994	1,887	72	76	55
4	s, u	381,000	216	1995	1,826	52	71	37
5	r, s, u	116,000	196	1995	1,826	69	72	50
6	r, s, u	200,000	2,222	1995	2,104	68	72	49
7	r, s	70,000	128	1996	1,608	72	71	51
8	s, u	1,400,000	1,985	1997	1,705	70	65	46
9	u	132,000	7,333	1997	1,447	76	58	44
Means						60	71	47

*r = rural, s = suburban, u = urban.

cooperation rate by the mail-phase response rate). Representiveness of each sample was evaluated by comparing general demographic characteristics with U.S. Census estimates. In each community, survey respondents were more likely to be female, to have one to two years more educational attainment, to be one to three years older, and to have a household income slightly higher than that of the population. Racial/ethnic minority groups were typically under-represented as well. Community characteristics and survey response rates appear in Table 1. From the nine communities included in the database, a total of 17,506 cases are available for analysis.

Database Construction

For each community survey, data were entered by a commercial data entry service, verified by 20 percent re-entry, and analyzed using SPSS©. Each survey employed a similar data collection methodology and instrument allowing the survey databases to be combined to create one master database for this research and other purposes. In total, 170 data elements are similar across all of the surveys.

The communities represented in the database range from a single city, to multicounty areas cutting across state lines. The communities are a mix of rural, suburban, and urban areas with total populations ranging from approximately 41,000 to 1.4 million persons.

Analysis Approach

Two analyses were conducted for this study. First, to examine the usefulness of the SF-12 as a measure of community or population health status, mean PCS-12 and MCS-12 scores for each community were compared with four traditional health measures. As measures of community health status, infant mortality and homicides per 100,000 population were selected. These measures are a subset of the Healthy People 2000 Health Status Indicators identified by the Centers for Disease Control and Prevention. Although not one of the CDC consensus indicators, the age-adjusted all-cause mortality rate was also selected in recognition of the effects of age on health status. The number of physicians per 1,000 population for each community was included as a comparison measure to incorporate the influence of access to medical care on health status. The choice of these measures was based on several factors including availability of the data and their frequency of use in typical community assessment activities. Data were obtained from the CDC's WONDER database for all-cause age-adjusted mortality and

homicide rates for the period 1992–1996 for each community. In addition, infant mortality and physician/population ratio data were obtained from the National Center for Health Statistics for 1993, the most current year available for all sites (National Center for Health Statistics, *Healthy People 2000 Statistical Notes*, vol. I, no. 1 [Fall 1991]). Also for those communities, Physical Components Summary (PCS-12) and Mental Components Summary (MCS-12) scores were calculated from the items that comprise the SF-12 functional health status instrument using the analysis program/recommended protocol provided by The Health Institute of The New England Medical Center using SPSS© software. SF-12 physical and mental components scores for each community were looked at using Pearson's correlations to examine the relationships between community health status indicators and SF-12 scores (Ware, Kosinski, and Keller 1996a,b).

The second analysis examined SF-12 scores in relation to measures of the social determinants of health for respondents from all nine communities. Although substantial discussion has taken place about the concept of social determinants of health, consensus does not exist about a specific model or set of measures, or their relationship to health status and each other. Patrick and Wickizer (1995) described social determinants of health and quality of life, with examples of "social class, gender inequalities, and racism" as factors that influence "how long" and "how well" people live. In their work, Patrick and Wickizer also recognize that social determinants of health operate at both the individual and group ecological levels. Amick, Levine, Tarlov and Walsh summarize the evolution of research and thinking about the social determinants and include factors such as class, race, education, gender, marital status, employment, and the quality of jobs. They also point out, however, that these factors are "deeply embedded in larger systems of social relations" and, therefore, that political, economic, and cultural factors must be considered in discussing social determinants of health (Amick et al. 1995).

We chose an examination of correlates of SF-12 component scores with social determinants of health variables from the community surveys as the first step in testing the hypothesis that the influence of the social determinants of health can be reflected in the overall physical and mental functioning of the population. In this analysis, population is assumed to be the aggregate of the individuals included in the database. This analysis, too, may be the starting point for discussion about the value of including functional health status measures as part of a community health status assessment.

One-way analysis of variance, followed by post hoc testing with the Scheffé multiple comparison procedure, was performed to identify significant

differences between the mean SF-12 scores of survey respondents that are grouped by selected demographic and social determinants of health variables. When we reject the null hypothesis that all of the group means are equal, this rejection indicates that at least one pair of means (maybe more) are unequal. The Scheffé post hoc procedure was used to determine which pairs of means are significantly different by testing all possible combinations. This approach is more conservative than other tests, requiring a larger difference for significance, and it is appropriate with larger sample sizes. The .05 significance level was used. The variables examined in this analysis for their relationships with functional health status measures were gender, race/ethnicity, total household income, age, educational attainment, and employment status.

RESULTS

Comparison of the SF-12 Component Scales with Traditional Community Health Measures

To determine the adequacy of the SF-12 components scores as overall measures of community health status, Pearson's correlations between mean PCS-12 and MCS-12 scores for each of the nine communities, and the selected community health status measures, were calculated (Table 2). The correlation between PCS-12 and MCS-12 scores (0.488) is moderate and expected, but not significant, supporting the argument that they represent related but different aspects of health status. No correlation is found between infant mortality and physical functioning (−0.160) or mental functioning (−0.038). The correlation between PCS-12 scores and the age-adjusted mortality rate was significant (−0.672), while the correlation with the number of physicians per 1,000 population rate was substantial (0.473) but not significant. No correlation was found between PCS-12 scores and homicide rates. MCS-12 scores and age-adjusted mortality did not correlate; however, the homicide rate was significantly correlated with mental health functioning (0.671). A moderate but not significant correlation was found between infant mortality and the physicians to population measure (0.528).

Based on these results we can say that the two SF-12 subscales for physical and mental health functioning correlate with two other "traditional" measures of community health: physical health functioning with age-adjusted mortality and mental health functioning with homicide rates. Among the "traditional" measures—infant mortality, all-cause age-adjusted mortality, homicide rate, and physicians per 1,000 population—no correlations were found

Table 2: Correlations of Community Health Status Measures

	N = 9	*MCS-12 Score*	*Infant Mortality*†	*All-Cause Age-adjusted Mortality Rate*‡	*Homicide Rate*§	*Physicians per 1,000 Population*
PCS- 12 score	Pearson Corr.	0.488	−0.160	−0.672*	−.041	0.473
	(Sig.)	(.182)	(.681)	(.047)	(.916)	(.198)
MCS-12 score			−0.038	−0.297	.671*	0.343
			(.923)	(.438)	(.048)	(.366)
Infant mortality				−0.068	.339	0.528
				(.861)	(.372)	(.144)
Age-adjusted mortality rate					.232	−0.310
					(.549)	(.416)
Homicide rate						−0.428
						(.250)

*Significant at the .05 level.

†Per 1,000 live births.

‡Per 100,000 population, 1940 method.

§ Rate/100,000 population.

with each other. In addition, the physician to population ratio approached a significant correlation with the physical health functioning scale and infant mortality rates.

We believe that the SF-12 mental and physical health scales can be meaningfully used in examining health status of communities and populations based on correlations with community health status measures that describe whole population mortality and resources (such as physician to population ratio) but not with subpopulation-specific measures, such as infant mortality.

Comparison of SF-12 Component Scores with Social Determinants Measures

Before examining the social determinants measures and their relationships to functional health status scores, those scores were examined for each calendar year in which surveys were conducted to determine if differences that might be found could result from the passage of time. Although there were significant differences by year for both PCS-12 and MCS-12 scores ($p = .000$ for each), Scheffé post hoc analysis (not shown) leads to the conclusion that survey year and health status scores were not meaningfully related.

The first social determinant measure examined was gender. Significant differences by gender were found for both MCS-12 and PCS-12 scores ($p = .000$ and $p = .001$, respectively). The mean scores for females were lower than males for both the PCS-12 and MCS-12 (females 49.08 versus 49.68 for males, and females 47.64 versus 48.15 for males: PCS-12 and MCS-12 scores, respectively). This result is similar to that reported in other literature (Ware, Bayliss, Rogers, et al. 1996; Ware et al. 1993).

Differences on the basis of race/ethnicity were found for the MCS-12 ($p = .000$) but not for PCS-12 scores ($p = .339$). MCS-12 scores (see Table 3), were grouped by Scheffé analysis into two subsets. Latino/Hispanic and White were unique to the lower mean scores category, while Black was the only unique category in the higher mean subset; the All Others category was found in both subsets. These racial/ethnic categories were created out of the categories available from the various community surveys.

The next social determinant examined was total household income. Three income categories were created by comparing total household income against the values for the Federal Poverty Level (FPL) on the basis of the total number of persons in the household and the year the survey was conducted: Poverty was defined as 0–100 percent of the FPL, Low Income as 101 percent to 185 percent of the FPL, and Above Low Income as 186 percent of the FPL and higher. (These cutoff points were selected because they are frequently

Table 3: Mean MCS-12 Scores by Race/Ethnicity

		Scheffé MCS-12	
Race/Ethnicity	N	*1*	*2*
Latino/Hispanic	555	47.22	
White	13,136	47.72	
All others	509	48.11	48.11
Black	887		49.10

Table 4: PCS-12 and MCS-12 Scores by Income

		Scheffé PCS-12	
Income	N	*1*	*2*
Low Income	1,933	47.97	
Poverty	1,571	48.04	
Above Low	11,175		49.72
		Scheffé MCS-12	
Poverty	1,571	47.09	
Low Income	1,933	47.28	
Above Low	11,175		47.99

used by a variety of health and human services programs to determine eligibility.) Significant differences were found for both PCS-12 and MCS-12 scores on the basis of income category ($p = .000$ for each). As Table 4 shows, two subsets were found for each functional health status measure. Poverty and Low Income appear as one subset, while Above Low Income is the second subset for both PCS-12 and MCS-12 scores.

Although related to income, employment status was examined separately. Respondents were categorized as Employed Full-time, Employed Part-time, Full-time Homemaker, Student, Retired, or Unemployed. PCS-12 and MCS-12 scores were significantly different within employment categories ($p = .000$ for each), and different groupings of employment categories were found in Scheffé post hoc analysis for the two health status measures, as shown in Table 5. For the PCS-12, three subsets were found. The unique means were Retired in the lowest mean subset, Unemployed and Full-time Homemaker in the middle subset, and Employed Full-time in the subset with the highest mean value. For MCS-12 scores only two subsets were found. Student and

Unemployed were found only in the first subset, while Retired was found only in the second subset.

Significant differences were found on the basis of six age categories (18–34, 35–44, 45–54, 55–64, 65–74, and 75 and older) between PCS-12 and MCS-12 scores ($p = .000$ for both). Scheffé post hoc testing found three subsets of age groups for PCS-12 scores and two subsets for MCS-12 scores (Table 6). Persons ages 65–75 and 75+ were unique to the lowest mean PCS-12 scores group, the 55–64-year-old group was unique to the middle subset, and persons ages 18–34 and 35–44 were unique to the highest mean PCS-12 scores subset. For MCS-12 scores, the lowest mean scores group contained the 18–34-year-old group, with all other ages in the higher mean subset. The ordering of the age groups in each subset was generally linear: from oldest to youngest for PCS-12, and youngest to oldest for MCS-12.

Educational attainment was the final social determinant examined. Significant differences were found for both PCS-12 and MCS-12 scores ($p = .000$ each). As can be seen in Table 7, three subsets were found for PCS-12 scores: (1) 0–12 years of education, (2) 12–16 years of education, and (3) 17 or more years of education. MCS-12 scores were clustered into two subsets: (1) 0–12 years of education and 12–16 years of education, and (2) 17 or more years of education.

Table 5: PCS-12 and MCS-12 Scores by Employment Status

		Scheffé PCS-12		
Employment Status	N	*1*	*2*	*3*
Retired	2,409	47.11		
Unemployed	472		48.84	
Full-time homemaker	1,392		49.02	
Student	475		49.93	49.93
Part-time	1,869		50.16	50.16
Full-time	6,866			50.78
		Scheffé MCS-12		
Student	461	47.24		
Unemployed	472	47.28		
Part-time	1,829	47.60	47.60	
Full-time homemaker	1,344	47.75	47.75	
Full-time	6,779	47.89	47.89	
Retired	2,409		48.88	

Table 6: PCS-12 and MCS-12 Scores by Age Group

		Scheffé PCS-12		
Age Group	N	*1*	*2*	*3*
75 and older	782	46.70		
65–74	1,604	47.53		
55–64	1,719		48.64	
45–54	2,450		49.65	49.65
18–34	4,633			49.93
35–44	3,171			50.74
		Scheffé MCS-12		
18–34	4,633	46.88		
35–44	3,171		48.10	
45–54	2,450		48.34	
75 and older	782		48.75	
55–64	1,729		48.89	
65–74	1,604		48.93	

Table 7: PCS-12 and MCS-12 Scores by Educational Attainment

		Scheffé PCS-12		
Educational Attainment	N	*1*	*2*	*3*
0–12 years	1,776	46.70		
13–16 years	5,538		48.69	
17 or more years	7,793			50.33
		Scheffé MCS-12		
0–12 years	1,776	47.29		
13–16 years	5,538	47.31		
17 or more years	7,793		48.28	

DISCUSSION

The SF-12 has largely been used in clinical environments to measure the functional health status of individual patients or to measure the health status of health plan enrollees. Based on the results of our analysis, which correlated

SF-12 component scores to other measures of community health status, and the social determinants of health within a community setting, the instrument appears to have value if it is included as one component of a comprehensive community or population health status assessment.

The SF-12 can be used to measure the *current* health status of a population or community. The value of a simple, real-time tool for describing the functional health status of communities or populations—a tool that can be linked with risk factor and disease prevalence as well as with utilization and social determinants variables—is substantial, particularly in the planning, implementation, and evaluation of health status improvement initiatives. Although measures such as infant mortality and age-adjusted mortality are of great value in documenting the end result of health-related processes, they are of limited value in more proximate considerations of factors influencing the health of populations or communities. These proximate considerations involve the social determinants of health.

Figure 1 is a summary of the significant relationships between the social determinants in this study and physical and mental health functioning. For the purposes of discussion, in this study physical component scores and mental

Figure 1: Overall Relationships of Social Determinants to Physical and Mental Functioning Scores

Social Determinants	Physical Components Score	Mental Components Score
Gender	Female = Low	Female = Low
Race/Ethnicity	No differences	Latino/Hispanic and White = Low Black = High
Income	Poverty and Low Income = Low Above Low Income = High	Poverty and Low Income = Low Above Low Income = High
Employment Status	Retired = Low Full-time Homemaker and Unemployed = Medium Full-time Employed = High	Student and Unemployed = Low Retired = High
Age Group	65–74 and 75+ = Low 18–34 / 35–44 = High	18–34 = Low 35–44 / 45–54 / 55–64 / 65–74 / 75+ = High
Education	0–12 years = Low 13–16 years = Medium 17 or more years = High	0–12 years and 13–16 years = Low 17 or more years = High

component scores were described as "low," "medium," and "high" in their relationship to social determinants based on Scheffé groupings. When we examine measures of social determinants of health as they relate to functional health status measures, two different patterns emerge. Low PCS-12 scores are found for persons with lower incomes, retired persons, older persons, persons with less educational attainment, and widowed persons. Although further examination is warranted, the authors feel safe in suggesting that age is driving the relationship between these factors and physical functioning. The pattern for MCS-12 scores, however, is less clear. As with physical functioning scores, lower income and lower educational attainment are associated with lower MCS-12 scores; however, in the latter two social determinants, the relationships between employment status, age, and MCS-12 scores are reversed. These relationships raise a series of questions:

1. If age is the driving factor determining physical functioning, is it necessary to measure and/or attempt to intervene at the level of other social determinants of health in an effort to improve or reduce the decline of physical functioning?
2. Do lower MCS-12 scores result from unemployment, lower income, and less education, or does lower mental health functioning produce these factors?
3. Why is the relationship between the youngest age group (18–34) and lower mental health scores so strong (in comparison to all other age groups)?
4. Does a standardized measure adequately address differences in definitions, perceptions, attitudes, and values related to mental health functioning found between different racial/ethnic groups?

The database of community survey respondents examined in this analysis could also be used to ask and answer the question, What is the minimum set of factors required to adequately represent the influence of social determinants on the health status of a population or community? In answering this question, two analytical approaches can be pursued by using the database compiled for this exploratory analysis: (1) examination of these data using the community as the unit of analysis (as opposed to the individual as the unit of analysis aggregated by social determinants factors in this analysis), and (2) exploration of the interactions between the social determinant variables (e.g., to understand the extent to which age is the underlying factor that determines physical components scores).

A related question to consider is that of identifying ways to test scores when examining community or population health data. In this study, we have taken the straightforward approach of testing means to determine relationships between factors. However, it can be effectively argued that the "average" is not what is of interest in identifying or attempting to intervene with factors influencing community health status, but rather the "variability" that should be of greater concern. Ware et al. have demonstrated an approach for use of the SF-12 in this manner to examine differences in health status of a Medicare population (Ware, Kosinski, and Keller 1996; Ware, Bayliss, Rogers, et al. 1996). An additional question for further research is this: Would examination of the distribution or variability of SF-12 scores rather than the means lead to different conclusions?

This discussion would be incomplete without mentioning the issue of interpretation of SF-12 scores. Designed to approach a mean of 50 with a standard deviation of 10 in a representative sample of the U.S. population, these scores vary by approximately one to three points between ten-year age groups. In a clinical context, attention focuses on differences in "clinical state" that one can determine by differences in functional health status scores. In community or population application, however, means to describe differences, such as those found in this article, are less clear; they may or may not be "meaningfully" different despite their statistical significance. Because of the exploratory nature of this study no effort was made to standardize interpretation; but further examination of these and comparable data should be pursued so that meaningful differences in functional health status, in a community or population context, can be described and similarly interpreted.

Finally, in considering the use of the SF-12 as a single measure or component of an array of other measures of community health status, it can be argued that other approaches to measuring community health status, such as the "environmental measures" approach of Cheadle et al., warrant examination as alternatives to the community-based survey methodologies required by this approach. Even though no one can disagree on the differences in relative cost of obtaining these data, the pertinent question—and one beyond the scope of this analysis—is that of the value of the resulting information (Cheadle, Wagner, Koepsell, et al. 1992).

Limitations

The methodology used to collect these data imposes limitations and requires caution in interpreting the results. The survey approach includes biases against persons without telephones or mailing addresses, and against those

who do not read English or Spanish at the seventh grade level or who may feel threatened by health-related survey or interview activities. The net effect in individual communities, and in the aggregate, is that this approach underrepresents the homeless, those with the lowest incomes, and/or non-English-speaking or low-literacy individuals. Further, the communities themselves were not randomly selected: despite the diverse environments they represent, selection bias must certainly be considered.

CONCLUSIONS

The SF-12, together with its physical and mental components scores, has been used chiefly in clinical environments for disease-specific health status monitoring. In a series of community health status assessments we used the instrument to measure health status in general population samples from multiple communities. Based on the results of the analysis reported here, the SF-12 and its component scales appear to be valid and useful tools to use in identifying differences in a population on the basis of social determinants of health. Its value, in part, is that it provides a mechanism for increasing our ability to identify factors that influence community-level health status so that we might be more effective in planning ways to improve health status and in documenting changes in health status based on improvement strategies.

After validating the community-level functional health status scores against traditional community health status measures, the balance of this article reported the results of individual respondent data, grouped on the basis of social determinant characteristics. The data were collected in random samples of the general population at the community level ($n = 9$); therefore, analysis of the social determinants data and their relationships to PCS-12 and MCS-12 scores is also possible at that level. Examination of these data on the basis of community variation in social determinants and the relationship of community social variation with health status would provide additional insights useful in planning community-level health status improvement interventions.

The results of this analysis suggest that the SF-12, and its component scores, could serve as an appropriate dependent variable in a subsequent analysis of these or similar data to explore the nature of the relationships between the social determinants of health and health status, with the objective of refining a set of measures of the social determinants that is both parsimonious and of practical value in efforts to improve the health status of community populations.

The historical use of the SF-12 in testing clinical improvement also suggests that it could serve as an instrumental bridge between the measurement of health status in a clinical setting and the measurement of health at a population or community level. As the need to understand the factors influencing the health status of communities and populations continues to build, driven by the forces of financing, service delivery planning, and regulation, the SF-12 could contribute substantially to meeting the data and analysis needs of the future, particularly in view of its potential linkage from populations to patient and clinical outcomes data. This linkage would serve as a powerful tool in expanding the ability to plan and evaluate efforts to improve population health across the patient-population "bridge." It warrants additional consideration.

ACKNOWLEDGMENTS

Dozens of communities and institutions have undertaken the task of measuring the health status of their populations. Those contributing to this database have taken an additional step by sharing the results of their work with others (in this and other contexts). We wish to acknowledge their hard work and hope that this article further encourages them in their endeavors to improve the health status of their communities.

Additionally, colleagues at The Health Institute, New England Medical Center and Harvard University School of Public Health, including Drs. Al Tarlov, Dana Safran, Bill Rogers, John Ware, and Mark Kosinski, among others, have encouraged us in our pursuit of excellence in the assessment of population health status and in the application of that information to health improvement efforts.

The authors would also like to acknowledge Dr. James Reed, Dr. Kenneth McLeroy, and Mr. Dan Callahan for their assistance in the collection, analysis, and preparation of these data and for their critiques of earlier manuscripts.

REFERENCES

Amick, B. A., S. Levine, A. R. Tarlov, and D. C. Walsh, eds. 1995. *Society and Health.* New York: Oxford University Press.

Blum, H. 1974. *Planning for Health: Development and Application of Social Change Theory.* New York: Human Science Press.

Cheadle, A., E. Wagner, T. Koepsell, A. Kristal, and D. Patrick. 1992. "Environmental Indicators: A Tool for Evaluating Community-based Health Promotion Programs." *American Journal of Preventive Medicine* 8 (2): 345–50.

Eden, J. 1998. "Measuring Access to Care Through Population-based Surveys: Where Are We Now?" In *Measuring Access to Care Through Population-based Surveys in a Managed Care Environment,* edited by A. B. Bindman and M. R. Gold. *Health Services Research* 33 (3, Part II): 685–707.

Evans, R. G. 1994. *Why are Some People Healthy and Others Not? The Determinants of Health of Populations.* New York: Aldine de Gruyter.

Felix, M., and J. N. Burdine. 1995. "Taking the Pulse of the Community." *Healthcare Executive* 10 (4): 8–10.

Gold, M., and B. Stevens. 1998. "Synopsis and Priorities for Future Efforts." In *Measuring Access to Care Through Population-based Surveys in a Managed Care Environment,* edited by A. B. Bindman and M. R. Gold. *Health Services Research* 33 (3, Part II): 611–24.

Illich, I. 1976. *Medical Nemesis: The Expropriation of Health.* Toronto, ON: Bantam Books.

Institute of Medicine. 1998. *The Future of Public Health.* Washington, DC: National Academy Press.

Joint Commission for the Accreditation of Healthcare Organizations. 1997. *Comprehensive Accreditation Manual for Hospitals: The Official Handbook.* Chicago: Joint Commission.

Kaplan, G. A., and J. W. Lynch. 1997. (Editorial) "Whither Studies on the Socioeconomic Foundations for Population Health?" *American Journal of Public Health* 87 (12): 1409–11.

Lalonde, M. 1975. *A New Perspective on the Health of Canadians.* Ottawa, ON: Information Canada.

Marmot, M. G., M. Bobak, and G. D. Smith. 1995. "Explanations for Social Inequalities in Health." In *Society and Health,* edited by B. C. Amick, S. Levine, A. R. Tarlov, and D. C. Walsh, pp. 46–92. New York: Oxford University Press, Inc.

McHorney C. A., M. Kosinski, and J. E. Ware. 1994. "Comparisons of the Costs and Quality of Norms for the SF-36 Health Survey Collected by Mail Versus Telephone Interview: Results From a National Survey." *Medical Care* 32 (6): 551–67.

Medical Outcomes Trust. 1995. "Community-Based Health Measurement." *Medical Outcomes Trust Bulletin* 3 (1): 1–3.

Patrick, D. L., and T. M. Wickizer. 1995. "Community and Health." In *Society and Health,* edited by B. C. Amick, S. Levine, A. R. Tarlov, and D. C. Walsh, pp. 46–92. New York: Oxford University Press, Inc.

Steckler, A. B., J. P. Allegrante, D. Altman, R. Brown, J. N. Burdine, R. M. Goodman, and C. Jorgensen. 1995. "Health Education Intervention Strategies: Recommendations for Future Research." *Health Education Quarterly* 22 (3): 307–28.

Tarlov, A. R., J. E. Ware, S. Greenfield, E. C. Nelson, E. Perrin, and M. Zubkoff. 1989. "The Medical Outcomes Study: An Application of Methods for Monitoring the Results of Medical Care." *Journal of the American Medical Association* 262 (7): 925–30.

Ware, J. E., K. K. Snow, M. A. Kosinski, and B. Gandek. 1993. *SF-36 Health Survey Manual and Interpretation Guide.* Boston: The Health Institute, New England Medical Center.

Ware, J. E., M. S. Bayliss, W. H. Rogers, M. Kosinski, and A. R. Tarlov. 1996. "Differences in 4-Year Health Outcomes for Elderly and Poor, Chronically Ill Patients Treated in HMO and Fee-for-Service Systems: Results from the Medical Outcomes Study." *Journal of the American Medical Association* 276 (13): 1039–47.

Ware J. E., M. Kosinski, and S. D. Keller. 1996a. *SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales.* Boston: The Health Institute, New England Medical Center.

———. 1996b. "A 12-Item Short-Form Health Survey (SF-12): Construction of Scales and Preliminary Tests of Reliability." *Medical Care* 32 (3): 220–33.

Wilkinson, R. G. 1986. *Class and Health.* London: Tavistock Publications.

William, R. M. 1990. "Rx: Social Reconnaissance." *Foundation News* 31 (4): 24–29.