Utility of Health Data from Social Surveys: Is There a Gold Standard for Measuring Morbidity?
Author(s): Kenneth F. Ferraro and Melissa M. Farmer
Source: *American Sociological Review*, Vol. 64, No. 2 (Apr., 1999), pp. 303-315
Published by: American Sociological Association
Stable URL: http://www.jstor.org/stable/2657534
Accessed: 11/10/2011 15:08

UTILITY OF HEALTH DATA FROM SOCIAL SURVEYS: IS THERE A GOLD STANDARD FOR MEASURING MORBIDITY?*

Kenneth F. Ferraro
Purdue University

Melissa M. Farmer
University of California at Los Angeles

Most sociological and epidemiological studies of health status in adulthood rely on reports of morbidity from respondents to social surveys. This study compares self-reported morbidity with indicators of morbidity from physicians' evaluations and examines the predictive validity of each indicator on self-assessed health and mortality in adulthood. Special attention is given to differences in the measures between white and African American adults. Adults from a large national survey received a detailed medical examination by a physician; they also were asked about the presence of 36 health conditions. Results indicate that self-reported morbidity is equal or superior to physician-evaluated morbidity in a prognostic sense. Both types of morbidity predict self-assessed health for white respondents, but physician-evaluated morbidity is not related to either self-assessed health or mortality for African American respondents.

Reliance on survey research methods by social scientists in the past half century, coupled with the widespread availability of data archives, means that much empirical knowledge in the social sciences depends on survey methods. In many situations, survey methods are the data collection strategy of choice, especially when coupled with appropriate sampling procedures. For studying quality-of-life indicators, attitudes, or characteristics of the self, survey methods may provide the most useful data because respondents are the best source of information for phenomena that typically "have no objective external referent" (Angel and Gronfein 1988:464). On the other hand, for researching topics such as church attendance or drug or alcohol consumption, survey methods may not be optimal because what respondents report may not accurately reflect objective realities (Bradburn, Rips, and Shevell 1987; Hadaway, Marler, and Chaves 1993). While relatively few investigators have examined how respondent self-reports of factual information correlate with other information sources, evidence is growing to show that respondent self-reports are influenced by contextual clues and respondent motivation (Nadeau and Niemi 1995; Presser and Traugott 1992).

The accuracy of survey respondents' reports of objective conditions affects virtually all fields of sociology. This paper examines this phenomenon with health data from physicians and survey respondents. Our analysis focuses on morbidity, a phenomenon that has an objective external referent. Some symptoms of disease may be difficult to corroborate from other sources, but the application of a diagnosis to a set of symptoms has a clear reference, largely owing to pathology tests and the perceived legitimacy of diagnoses from physicians or other health-care

*Direct all correspondence to Kenneth F. Ferraro, Professor of Sociology, Stone Hall, Purdue University, West Lafayette, IN 47907-1365 (ferraro@purdue.edu). This paper was presented at the 1996 annual meeting of the Gerontological Society of America in Los Angeles. Support for this research was provided by a grant from the National Institute on Aging to the first author (AG11705). We appreciate help from Ronald Angel and Ellen Idler with the construction of selected morbidity variables, and Jessica Kelley-Moore and Jamie DeCoster with statistical computation. Ralph Cherry, Daniel Clark, Jean Findlay, Neal Krause, George McCabe, and three anonymous *ASR* reviewers provided helpful comments on earlier versions of the paper. Data were made available by the Inter-university Consortium for Political and Social Research, Ann Arbor, MI. Neither the collector of the original data nor the Consortium bears any responsibility for the analyses or interpretations presented here.

practitioners. Medical knowledge is typically seen as the ultimate information source about the existence of disease, and physicians are the widely acknowledged arbiters of the presence of disease in an individual. In a sense then, the profession of medicine is "engaged in *the creation of illness as a social state which a human being may assume*" (Freidson 1970: 205, emphasis in original).

Although data from physicians are usually regarded as the "gold standard" for assessing health status, this research compares self-reports of morbidity from survey respondents with evaluations of respondents' morbidity by physicians. Previous studies have compared data from these two sources, but most focus on the detection of bias or recall error for self-reported morbidity with physician-evaluated morbidity seen as true prevalence (Coughlin 1990). This study poses a different question: Which data source is most useful in a prognostic sense for understanding health trajectories? This question is addressed by examining the predictive validity of *self-reported* and *physician-evaluated morbidity* in relation to two outcomes: self-assessments of health and mortality. A second aim is to determine if these relationships differ for white and African American respondents. Because much of what is known about the prevalence of illness and the extent of health inequality depends on respondents' reports of illness, we examine whether black respondents' and white respondents' reports of morbidity concur with physicians' evaluations or are systematically different.

HEALTH INFORMATION FROM SURVEY RESEARCH

Many studies have compared respondent reports and physician evaluations of morbidity: Most such studies are limited to investigations of single diseases, such as arthritis (Altman 1991) or heart disease (Smith et al. 1993), or to a set of conditions like neurological disorders (Anderson, Schoenberg, and Haerer 1988). Some of these studies do not ask respondents whether they have a medical condition, but simply examine the relationships between medical measures, such as electrocardiogram tracings, and respondent reports of symptoms (Smith et al. 1993).

Typically, studies comparing self-reported morbidity with information from physicians use medical records. In an extensive review, Harlow and Linet (1989) find considerable evidence for the accuracy of respondent recall on morbidity when comparing questionnaire data and medical records. However, other studies show that this agreement is limited to selected diseases, especially diseases that have clear and unambiguous diagnostic criteria (Colditz et al. 1986; National Center for Health Statistics 1965).

A major limitation of most studies comparing self-reported morbidity with medical records is that only *patients* are considered. Thus, estimates of the association between self-reported and physician-evaluated morbidity from these studies are influenced by the use of medical care. Interpretations from medical records also depend on temporal decisions that define who is in the patient group. For example, using records from patients who have recently been to a physician or hospital will focus disproportionately on more frequent users of medical care. Fortunately, several studies avoid this problem by using surveys of population samples to compare self-reported morbidity with data gathered from medical tests (Ford et al. 1990), physician interviews (Pecoraro et al. 1979), physician evaluations (Heliövaara et al. 1993), or some combination of all three (Guralnik et al. 1995).

The last two studies mentioned above are especially relevant because they entail physician evaluations of more than one disease. Heliövaara et al. (1993) compare morbidity reported by over 7,000 adult Finnish respondents with that discovered by physicians in an independent medical examination. They find only moderate agreement on most conditions, but substantial agreement on comparisons for cardiovascular disease. More recently, data from the ongoing Women's Health and Aging Study also show moderate agreement (Guralnik et al. 1995; Guralnik et al. 1996). Of the 17 diseases considered, the authors find that self-report data are consistent with physician evaluations of morbidity for some conditions (e.g., stroke, cancer), but not for others (e.g., angina, arthritis).

Taken together, studies of the association between self-reported morbidity and morbidity data gathered by physicians show modest agreement. The strength of the association varies across studies, and across diseases for those studies that consider more than one disease. For all studies, "true prevalence" is defined by data from physicians. Physicians' information is treated as the "gold standard," and reports that are inconsistent with physician reports are considered false positive or false negative cases. When data from respondents and physicians do not agree, the presumption is that respondents are underreporting or overreporting medical conditions (Angel and Guarnaccia 1989; Heliövaara et al. 1993). In the final analysis, however, biopsy or autopsy may be the gold standard, for any physician evaluation of morbidity can be revised by the findings of a biopsy or autopsy.

HEALTH DATA AMONG ETHNIC GROUPS

The substantial health inequality in American society by racial and ethnic groups suggests that members of minority groups should report greater morbidity and poorer health. Indeed, dozens of studies of morbidity, disability, and health ratings over the life course reveal that African Americans and Hispanic Americans have poorer health than do non-Hispanic whites (American Medical Association 1991; Ferraro and Farmer 1996; Hummer 1996). There is little debate that most African Americans, Hispanic Americans, and Native Americans are health disadvantaged, but the extent of the inequality is a matter of continuing interest.

Because most ethnic minorities have less frequent or delayed contacts with physicians (Himmelstein and Woolhandler 1995; Wolinsky et al. 1989), they also may be less likely to have a definitive diagnosis about conditions that bother them (and thus present to a physician at a more advanced illness stage). Moreover, when a black or Hispanic person is examined by a physician, he or she is also likely to be in an ethnically mixed dyadic relationship that may hinder the full flow of health communication. Indeed, Hall, Roter, and Katz (1988) found that black patients and Hispanic patients receive less information and positive talk. Thus, differences in contact and communication may result in an underestimate of the true prevalence of morbidity for some racial/ethnic groups and have dire consequences for the health of group members. Compared with physician visits, hospitalization rates are less discrepant across ethnic minority groups (Blustein and Weitzman 1995). This suggests that serious or life-threatening conditions (typically requiring hospitalization) may be more likely than nonserious chronic conditions to be detected. In short, while all medical conditions are probably underestimated for ethnic minorities, the bias is probably less for serious illnesses than for nonserious chronic conditions.

The purpose of the present research is to examine the predictive validity of physician-evaluated and self-reported morbidity among white respondents and African American respondents. Scores of studies over four decades have debated the merits of health data obtained from respondents and physicians, but we are unaware of any studies that have compared the two sources of data in a prognostic sense. Two basic research questions are addressed here:

(1) *Do physician-evaluated morbidity and self-reported morbidity manifest predictive validity on self-assessed health status and mortality? If both are useful in a prognostic sense, which manifests greater utility?*

Based on the extant literature, we hypothesize that self-reported morbidity is more predictive of self-assessed health status and that physician-evaluated morbidity is more predictive of survival.

(2) *Is the predictive validity of the two types of morbidity contingent on race?*

Because racial and ethnic factors influence the diagnosis and reporting of morbidity, we anticipate that the association between physician-evaluated and self-reported morbidity will be less among black respondents than among white respondents.

METHOD

Sample

Data for this analysis are from the National Health and Nutrition Examination Survey I

(NHANES-I) which collected medical examination and survey interview data in the period 1971–1975 (National Center for Health Statistics 1979). The survey is longitudinal, but the medical examination data were collected only during the baseline interview; therefore, change over time in the two types of morbidity cannot be examined. Our strategy is to examine the predictive validity of the two types of morbidity for self-assessed health at the baseline interview and for mortality over a 15-year observation period.

The sampling design was a multistage, stratified probability sample of noninstitutionalized persons ages 25 through 74, resulting in a total of 14,407 respondents. The analyses make use of the NHANES-I subsample, designed as a nationally representative sample, which was administered the "detailed component" in separate sessions including an extensive medical examination and phlebotomy (N = 6,913). Response rates were 98.6 percent for the interview and 70 percent for the medical examination (National Center for Health Statistics et al. 1987). The response rate for the medical examination was similar for black and white respondents, 70.2 percent and 69.4 percent, respectively.

Interviewers for the baseline survey coded race by observation as white, Negro, or other. (Race was requested only if the appropriate category could not be visually determined by the interviewer.) Follow-up surveys reveal that the sample included relatively few Asian Americans, Hispanic Americans, Native Americans, or persons in other ethnic categories. Because of their small numbers, it is not possible to analyze these groups in the present analysis. This study uses data from 5,968 white and 873 black respondents from the baseline survey and mortality data from the follow-ups of the panel completed in 1982 to 1984 and in 1987. The percentage of cases at baseline receiving the detailed component and traced through follow-ups is high (92 percent in 1982–1984, and 96 percent in 1987). Because NHANES-I is a stratified cluster sample, parameter estimates and significance tests were adjusted by Taylor linearization, which typically results in more appropriate tests of significance (Shah, Barnwell, and Bieler 1997).

Measures

NHANES representatives collected basic health and medical history information from sample members in their households and invited respondents to a nearby mobile examination center. The medical examination was extensive and included urinalysis and a blood panel. Most of the physicians conducting the general medical examination were white male internists within 10 years of completing medical residency. Also, they were commissioned officers of the Public Health Service and were trained by the NHANES staff on how to conduct the examination. Physicians were given the results of numerous NHANES laboratory tests and reviewed the medical history questionnaire for each respondent "on the day before the scheduled examination" (National Center for Health Statistics 1979:29). Thus, physicians had extensive information on each respondent before the examination, but no data were collected on how extensively this information was used.

Physicians evaluated morbidity in two ways. First, physicians examined eight anatomical areas in what is referred to as the *general medical examination*. The eight domains are: (1) head, eyes, ears, nose, and throat; (2) thyroid; (3) chest; (4) cardiovascular function; (5) abdomen; (6) musculoskeletal system; (7) neurological function; and (8) skin. The general medical examination (eight domains) was designed to uncover findings based on anatomy and function. Findings from each domain were binary coded (0, 1) and then summed.

Second, physicians conducted a *detailed medical examination* and recorded findings (abnormalities) according to the *International Classification of Diseases* (ICD) (World Health Organization 1967, rev. 8). The examination included detailed auscultation of the heart and a musculoskeletal examination, including range-of-motion and straight-leg-raising tests. Guided by laboratory tests, medical history, and the physical examination, the physician coded morbidity using 15 ICD domains: (1) infectious and parasitic diseases; (2) cancers and neoplasms; (3) endocrine, nutritional, and metabolic disorders; (4) diseases of the blood and blood-forming organs; (5) mental disorders;

diseases of the (6) nervous, (7) circulatory, (8) respiratory, (9) digestive, and (10) genitourinary systems; (11) diseases of the skin and subcutaneous tissue; (12) diseases of the musculoskeletal system and connective tissue; (13) congenital anomalies; (14) symptoms and ill-defined conditions; and (15) accidents, poisonings, and violence. Each ICD domain was coded as a binary variable (0 = no findings; 1 = one or more finding). The 15 binary codes were then summed; the resulting measure is referred to as *ICD examination.*[1]

Although the two protocols represent different classification schemes, each measures physician-evaluated morbidity. The simple correlation between the *general medical examination* and the *ICD examination* is .42.

Most of the remaining health measures used in the analysis were collected by NHANES staff in an interview with the respondent (e.g., medical history, general well-being). *Self-reported morbidity* was derived from a checklist question designed to identify which illnesses respondents had. Respondents were asked: "Has a doctor ever told you that you have . . . hypertension or high blood pressure?" (36 conditions were presented). Note that the answer is not a report about how the respondent feels about a specific condition but a report of a condition based on a medical encounter. Unlike some surveys that ask if a person has a particular condition, NHANES hinged the question on evaluation by a physician. Each condition was coded as a binary variable (1 = yes, 0 = condition not present). The conditions were then classified into those that are serious or life-threatening, and all remaining conditions (Ferraro and Farmer 1996). *Serious illnesses* include cancer, diabetes, heart failure (heart attack or heart trouble), hypertension, and stroke. Examples of the remaining conditions, hereafter referred to as *chronic illnesses*, include arthritis, asthma, bone fracture, cataracts, gout, psoriasis, and ulcer. The serious and chronic (nonserious) conditions were summed separately. The simple correlation between the two is .24.

Physical function at the baseline survey was assessed by asking about physical activity restrictions in both everyday activities and recreational activities. Respondents who answered that they were limited in both types of activity were assigned a score of 1 indicating "quite inactive" (0 otherwise).

Self-assessed health was measured with the question "Would you say that your health in general is excellent, very good, good, fair, or poor?" (Scores range from 1 = poor to 5 = excellent). Self-assessed health is a measure of the respondent's overall sense of health and is among the best predictors of medical care use and longevity (Angel and Gronfein 1988).

Vital status was determined at the follow-up surveys for all traced respondents, and death was confirmed by death certificates. Brief interviews were conducted with proxies of deceased respondents. Date of death was obtained for 1,337 decedents, so continuous-time event history models were applied. (Only all-cause mortality is considered; disease-specific cause of death from death certificates is not used.)

The independent variables span a broad range of factors known to be related to morbidity, self-assessed health, and mortality, either directly or indirectly (Hummer 1996; Mutchler and Burr 1991). These include indicators of health behavior such as obesity and smoking as well as enabling factors such as family income and type of medical insurance. Obesity was determined by the physician during the medical examination. Smoking was based on self-reports of consumption of cigarettes, cigars, and pipe tobacco at the time of the interview and during one's lifetime. The measurement of the remaining independent variables is straightforward.

FINDINGS

The analysis is organized in two stages. The first stage uses information from the NHANES-I survey to examine the relationships between the two types of morbidity and the predictive validity of each morbidity measure on self-assessed health. The second stage uses the follow-up surveys to compare the predictive ability of the measures of morbidity on mortality.

[1] ICD codes were designed to be comprehensive in reliably classifying many types of pathology and are widely used in health research (Idler and Angel 1990).

Table 1. Unstandardized Coefficients from the Regression of Morbidity and Self-Assessed Health on Selected Independent Variables: U.S. Adults, Ages 24 to 77, National Health and Nutrition Examination Survey, 1971–1975

	Physician-Evaluated Morbidity		Self-Reported Morbidity				Self-Assessed Health			
							Total			
Independent Variable	ICD Exam	General Medical Exam	Chronic Illness Model 1	Chronic Illness Model 2	Serious Illness Model 1	Serious Illness Model 2	Model 1	Model 2	Whites	Blacks
Age	.03***	.03***	.02***	.02***	.01***	.01***	–.004***	–.004***	–.005***	–.001
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Female	–.12***	.11***	.64***	.64***	.02	.02	.04	.03	–.05	–.07
	(.03)	(.03)	(.06)	(.06)	(.02)	(.02)	(.03)	(.03)	(.03)	(.10)
Black	.03	.05	–.43***	–.29**	.15***	.09*	–.24***	–.42***	—	—
	(.08)	(.08)	(.07)	(.09)	(.03)	(.04)	(.05)	(.07)	—	—
Lives alone	.05	–.12*	.00	.01	–.01	–.01	.20***	.20***	–.18***	.23
	(.05)	(.05)	(.07)	(.07)	(.03)	(.03)	(.05)	(.05)	(.05)	(.12)
Widow	–.02	.10	.06	.06	.04	.04	.19**	.19**	–.20**	.15
	(.07)	(.06)	(.10)	(.10)	(.04)	(.04)	(.06)	(.06)	(.07)	(.17)
Education (coded 0 to 7, with 7 = graduate school)	–.02	–.06***	.07***	.07***	–.01	–.01	.15***	.16***	.16***	.16***
	(.01)	(.01)	(.02)	(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.03)
Income (coded 1 to 12, with 12 = $25,000)	–.03**	–.02*	–.01	–.01	–.01***	–.01***	.07***	.06***	.07***	.04*
	(.01)	(.01)	(.01)	(.01)	(.01)	(.00)	(.01)	(.01)	(.01)	(.02)
Rural community	.01	.04	–.03	–.03	.00	.00	–.03	–.02	.01	–.16
	(.11)	(.08)	(.07)	(.07)	(.02)	(.02)	(.04)	(.04)	(.03)	(.10)
Medical insurance	–.15**	–.05	.03	.03	.01	.01	.10*	.10*	.09*	.12
	(.05)	(.05)	(.07)	(.07)	(.03)	(.03)	(.04)	(.04)	(.05)	(.10)
Medicaid	.05	.09	.18	.19	.12*	.11*	–.11	–.12	.07	–.25
	(.09)	(.10)	(.13)	(.13)	(.06)	(.06)	(.08)	(.08)	(.10)	(.15)
Regular physician	.02	.02	.15*	.15*	.09***	.09***	–.10*	–.10**	–.09*	–.18
	(.05)	(.05)	(.06)	(.06)	(.02)	(.02)	(.04)	(.04)	(.04)	(.11)
Obese	.42***	.22***	–.03	–.03	.14***	.14***	–.05	–.05	.06	.00
	(.06)	(.05)	(.07)	(.07)	(.03)	(.03)	(.03)	(.03)	(.04)	(.09)

(Continued on next page)

(Table continued from previous page)

Independent Variable	Physician-Evaluated Morbidity: ICD Exam	Physician-Evaluated Morbidity: General Medical Exam	Self-Reported Morbidity: Chronic Illness Model 1	Self-Reported Morbidity: Chronic Illness Model 2	Self-Reported Morbidity: Serious Illness Model 1	Self-Reported Morbidity: Serious Illness Model 2	Self-Assessed Health: Total Model 1	Self-Assessed Health: Total Model 2	Self-Assessed Health: Whites	Self-Assessed Health: Blacks
Smoker	.07*	.10***	.12*	.12*	–.03	–.02	–.05	–.04	.05	.05
	(.03)	(.03)	(.05)	(.05)	(.02)	(.02)	(.03)	(.03)	(.03)	(.09)
Past smoker	–.06	.03	.22***	.22***	.01	.01	–.02	–.02	.04	.21
	(.04)	(.04)	(.06)	(.06)	(.02)	(.02)	(.03)	(.03)	(.04)	(.13)
Restricted activity	—	—	—	—	—	—	–.54***	–.55***	–.56***	–.52***
							(.05)	(.05)	(.06)	(.12)
Physician-Evaluated Morbidity										
ICD examination	—	—	.24***	.25***	.07***	.07***	–.08***	–.09***	–.09***	.02
			(.04)	(.04)	(.01)	(.01)	(.01)	(.01)	(.01)	(.04)
General medical examination	—	—	.13***	.13***	.02**	.02**	–.04*	–.04*	–.05***	.02
			(.03)	(.03)	(.01)	(.01)	(.02)	(.02)	(.02)	(.04)
Self-Reported Morbidity										
Chronic illness	—	—	—	—	—	—	–.12***	–.12***	–.11***	–.16***
							(.01)	(.01)	(.01)	(.03)
Serious illness	—	—	—	—	—	—	–.30***	–.30***	–.30***	–.26***
							(.02)	(.02)	(.03)	(.06)
Black × ICD examination	—	—	—	–.13*	—	.05*	—	.17***	—	—
				(.06)		(.03)		(.04)		
Constant	.22	.81	.53	.67	.37	.32	2.47	2.29	2.52	2.30
R^2	.19	.18	.14	.14	.16	.16	.32	.32	.32	.28
Number of cases	6,801	6,801	6,801	6,801	6,801	6,801	6,794	6,794	5,929	865

Note: Numbers in parentheses are standard errors. Binary variables are coded 0 and 1.

$^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$ (two-tailed tests)

Columns 1 and 2 in Table 1 regress the two measures of physician-evaluated morbidity on the independent variables. Many relationships are similar—physician-evaluated morbidity, regardless of how it is measured, is higher among older persons, those with limited income, obese persons, and smokers. By contrast, several variables have different relationships with the two types of physician-evaluated morbidity. For instance, men scored higher on ICD morbidity while women scored higher on general medical morbidity, and education was significant only for the general medical examination. There were no differences in physician-evaluated morbidity between black respondents and white respondents. Higher ICD morbidity was also found among those without medical insurance.

Columns 3 through 6 use the self-reported morbidity measures as dependent variables and include physician-evaluated morbidity among the independent variables. Two models are presented for each type of self-reported morbidity. Model 1 is an additive specification. Based on the expectation that physician evaluations may be more likely to underestimate morbidity among black respondents, Model 2 adds an interaction term for black and ICD examination. This interaction term is significant for both chronic and serious illnesses (an interaction term for black and general medical examination also was tested but was not significant). Note that black respondents report lower levels of chronic illness, but higher levels of serious illness. Both forms of physician-evaluated morbidity are positively associated with self-reported chronic illness and serious illness. Moreover, the effect of ICD morbidity on both forms of self-reported morbidity is conditional by race (Model 2 in each instance): Among respondents with high ICD morbidity, African Americans were less likely than white respondents to report a chronic illness, but more likely to report a serious illness.

The remaining four equations (columns 7 through 10) specify self-assessed health as the dependent variable and include both types of morbidity. The results from Model 1 on the full sample (column 7) show that self-reported morbidity has a stronger negative effect than does physician-evaluated morbidity on self-assessed health (standardized coefficients, not shown, are –.18 for chronic illness, –.17 for serious illness, –.08 for ICD exam, and –.04 for general medical exam). Tests of the slope differences are also statistically significant ($p < .01$).[2] Additionally, older people, black respondents, and those with limited enabling resources have poorer self-assessed health. Model 2 (column 8) adds the interaction term for black and physician-evaluated ICD morbidity; both the interaction term and main effects are significant. Results show that the relationship between physician-evaluated ICD morbidity and self-assessed health is conditional by race. The black/white difference in slopes is substantial at low levels of physician-evaluated morbidity, but attenuates at higher levels of physician-evaluated morbidity.

The final two equations (columns 9 and 10) separate the white and black subsamples. All four indicators of morbidity are significant predictors of self-assessed health among white respondents, but neither indicator of physician-evaluated morbidity is related to self-assessed health among African American respondents. A test of slopes in each equation also reveals that self-reported serious illness is the strongest predictor ($p <$ .01). Given the different sizes of the white and black subsamples, supplementary analyses were undertaken to ensure that differences between the subsamples were not simply the result of the larger subsample of white respondents. The equation for the white subsample was reestimated on a random sample of 878 respondents to match the numbers of black respondents. The following variables were not significant in the reduced white subsample: lives alone, medical insurance, and regular physician. All four indicators of morbidity remained significant.[3]

[2] The test of the equality of regression coefficients was performed according to Gujarati (1988):

$$t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\mathrm{var}(\hat{\beta}_1) + \mathrm{var}(\hat{\beta}_2) - 2\,\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2)}}$$

where $\hat{\beta}_i$ refers to the respective unstandardized regression coefficients.

[3] Because African Americans make up just over 12 percent of the sample, differences in the results by race for the subsample analyses may be a result of the relatively small size of the black

Table 2. Unstandardized Coefficients from Proportional Hazards Models of Mortality: National Health and Nutrition Examination Survey, 1971–1987

Independent Variable	Total Sample		Whites		Blacks	
Age	.08***	(.00)	.08***	(.00)	.07***	(.01)
Female	–.61***	(.09)	–.62***	(.10)	–.72**	(.24)
Black	.07	(.10)	—		—	
Lives alone	.10	(.11)	.10	(.13)	.12	(.23)
Widowed	–.03	(.11)	–.06	(.14)	.32	(.25)
Education (coded 0 to 7, with 7 = graduate school)	–.04	(.03)	–.06	(.03)	.03	(.07)
Income (coded 1 to 12, with 12 = $25,000)	–.05**	(.01)	–.04*	(.02)	–.13**	(.04)
Rural community	–.06	(.08)	–.07	(.08)	.07	(.22)
Medical insurance	–.28**	(.10)	–.27*	(.12)	–.12	(.21)
Medicaid	.16	(.13)	–.07	(.17)	.82**	(.27)
Regular physician	–.01	(.10)	.05	(.12)	–.25	(.22)
Obese	.03	(.08)	.02	(.09)	.18	(.20)
Smoker	.56***	(.08)	.55***	(.09)	.69***	(.21)
Past smoker	.01	(.08)	.01	(.09)	–.01	(.30)
Restricted activity	.40***	(.10)	.47***	(.12)	–.04	(.27)
Physician-Evaluated Morbidity						
ICD examination	.00	(.03)	–.01	(.03)	.02	(.10)
General medical examination	.16***	(.03)	.17***	(.03)	.06	(.07)
Self-Reported Morbidity						
Chronic illness	–.04	(.02)	–.02	(.02)	–.17*	(.07)
Serious illness	.38***	(.05)	.37***	(.05)	.47***	(.13)
χ^2	1,613.48		1,413.04		233.41	
Degrees of freedom	19		18		18	
Number of cases	6,503		5,714		789	

Note: Numbers in parentheses are standard errors. Binary variables are coded 0 and 1. During the study period, 1,054 white respondents and 232 black respondents died.

*$p < .05$ **$p < .01$ ***$p < .001$ (two-tailed tests)

The second stage of the analysis considers how the two measures of morbidity influence mortality over the study period. Because date of death is known, event history analyses were performed to estimate these effects using Cox proportional hazards models. A total of 1,337 respondents died during the study period; of these, 1,286 are available for analysis once the predictor variables are considered. Results from the Cox

subsample. Reestimating the models on a random sample of the white respondents that equals the number of black respondents is one way to check for whether the differences for the two groups are due to the disparate number of cases. Some caution is warranted, however, when interpreting findings from the analysis of the reduced sample because of the unique covariance matrix. This is especially true for the mortality analyses that follow because of the racial differential in mortality. Given the lower mortality rate among whites, constraining the white subsample to equal the black subsample yields 158 white respondents who died during the survey period compared to 232 who died among the black subsample. Results from models of self-assessed health and mortality using the reduced white subsample show that some covariates are no longer significant, but the basic conclusions regarding the importance of the types of morbidity are unchanged.

models are presented in Table 2 for the total sample and by race.

The results for the total sample show that the strongest effects (reflected by *t*-values) are due to a higher mortality risk for older people, those reporting serious illnesses, smokers, and men. Although morbidity, measured both by physicians in the general medical examination and by self-reported serious illness, predicts mortality, self-reported serious illness is the stronger predictor ($t = -3.74$ in a test of slope differences). Respondents with limited income, no medical insurance, and those reporting restricted physical activity were also less likely to survive. Although reduced-form models show African Americans to have greater mortality risk, there were no racial differences in the final model shown in Table 2.

Findings for analysis of the white subsample resemble those for the full sample. For the black subsample, neither measure of physician-evaluated morbidity is a significant predictor of mortality, but both self-reported chronic illness and serious illness are significant. Serious illness is positively related to mortality, while chronic illness manifests a modest negative relationship to mortality—survivors face chronic disabling conditions. The relationships between mortality and age, female, income, and smoker parallel those for the full sample, but Medicaid is associated with higher mortality risk among African Americans. As before, the white subsample was reduced to the same size as the black subsample for supplementary analyses. The following variables were no longer significant: income, medical insurance, restricted activity, and general medical examination morbidity. Thus, by constraining sample sizes to be equal for the subsamples, self-reported serious illness is the only morbidity variable that is significant for both black respondents and white respondents.

Recognizing that alternative measures of morbidity may lead to different conclusions, we also compared physician-evaluated and self-reported morbidity based on coding from the *International Classification of Diseases* (World Health Organization 1967).[4] This was done to ensure that the results were not an artifact of coding and to determine if using an identical metric would modify the basic conclusions. The findings derived from these alternative specifications resemble those presented above when self-assessed health was the outcome measure. For the mortality analysis, the results are similar for the white subsample, but neither self-reported nor physician-evaluated morbidity based on the ICD codes is predictive among African Americans. Thus, neither set of analyses shows that physician-evaluated morbidity is a superior predictor of either self-assessed health or mortality.

DISCUSSION

Self-reported data are widely used in social science research, and comparing respondent reports with information from other sources sometimes leads to discrepant results. This research evaluated the utility of morbidity data reported by respondents with that provided by physicians for predicting health assessments and mortality. The two types of morbidity were only modestly correlated. All four indicators of morbidity—two reflecting physician-evaluated morbidity and two reflecting self-reported morbidity—were predictive of self-assessed health among white respondents. The relationships between self-assessed health and the two indicators of self-reported morbidity, however, were stronger than were those with physician-evaluated morbidity in the white subsample. For the black subsample, neither indicator of physician-evaluated morbidity was significant in predicting self-assessed health, but both indicators of self-reported morbidity

[4] While differences in the protocol and wording of the question for measuring physician-evaluated and self-reported morbidity preclude disease-by-disease comparisons, supplementary analyses were performed in which the self-reported illnesses were classified into the same 15 ICD codes used to measure physician-evaluated morbidity. The 15 dummy variables were then summed so that both self-report and physician-evaluated ICD morbidity could be compared using an identical metric. Despite the identical metric, the simple correlation between the two types of morbidity was modest (.28). It was not feasible to fold the information from the general medical examination into the physician-evaluated ICD codes, so analyses were completed with and without the general medical examination.

were significant. In summary, self-reported morbidity is the stronger predictor of self-assessed health among all respondents, and physician-evaluated morbidity is predictive of health assessments only among white respondents.

It is not surprising that self-reported morbidity is the stronger predictor of self-assessed health, given that the outcome was a subjective appraisal of health. Yet, even when mortality is the outcome, physician-evaluated morbidity was not the superior predictor. Self-reported serious illness and morbidity from the general medical examination predicted mortality among white respondents, but neither type of physician-evaluated morbidity was predictive of mortality among black respondents. In summary, the evidence shows that self-reported morbidity is equal or superior to physician-evaluated morbidity in a prognostic sense. Among white respondents, the difference in predictive ability of the two types of morbidity is modest, but among African American respondents, self-reported morbidity manifests stronger relationships.

These results suggest that self-reported data should not axiomatically be characterized as inferior solely because they come from respondents. The accuracy of survey data is an empirical question. The analyses presented here show that self-reported morbidity has considerable predictive validity for overall health and mortality. These findings bolster confidence in the use of such measures in the social and health sciences, but it may be judicious to avoid characterizing them as the "gold standard" until these findings are replicated and extended in future studies. The condition checklist in the NHANES-I was extensive; whether the shorter lists used in other surveys are equally useful is yet to be determined. These shorter lists emphasize serious illnesses, and these conditions should also have considerable predictive validity because, of the two indicators of self-reported morbidity, serious illness manifested stronger relationships than chronic illness with the outcomes under consideration. Still, future research should examine the predictive validity of alternative measures of morbidity.

To test an alternative procedure for measuring morbidity, we coded both physician-evaluated and self-reported morbidity according to the *International Classification of Diseases* (World Health Organization 1967), yielding morbidity variables with an identical metric. Although the basic conclusions were consistent using the alternative procedure, one disparate finding bears on the issue of using the measures of morbidity available in most health surveys. Self-reported serious illness was a significant predictor of mortality for white respondents and black respondents, but neither form of ICD-coded morbidity influenced mortality among black respondents. This finding suggests that categorizing self-reported illnesses into ICD codes may be useful in some contexts, but the procedure may not be optimal when morbidity is treated as an independent variable. ICD codes are useful for systematically classifying diseases, but the resulting 17 categories are fairly broad and based on body systems. Moreover, the 17 categories do not differentiate severity within each group (e.g., diseases of the circulatory system include both heart failure and heart murmur). The serious-illness and chronic-illness classification used here would treat heart failure as serious and heart murmur as chronic (nonserious). The distinction between serious and chronic conditions proved useful in a prognostic sense here, but merits further consideration in future research.

Social scientists should remain vigilant about issues of data quality, especially when self-reported information is used. Past research shows that deception is occasionally involved (Presser and Traugott 1992), but most of the time responses from survey participants "are likely to be biased by the assumptions that the respondents apply to the problem" (Bradburn et al. 1987:161). The type of information collected and the context of the questioning are also important when attempting to understand discrepancies between self-reported data and other information sources (Angel and Gronfein 1988; Nadeau and Niemi 1995). For instance, questions regarding sexually transmitted diseases probably contain more bias than would be the case for other conditions such as heart attack. Also, the NHANES-I form of the question relies on reports of whether a physician diagnosed a condition.

Self-reported morbidity prevalence may be higher when there is no reference to a physician's diagnosis.

Examining racial differences in the NHANES-I data clarifies the contextual bias of the morbidity measures used in many sociological and epidemiological studies. The correlations between the measures of morbidity from the physician examination and self-report varied less among white respondents (ranging from .21 to .26) than among black respondents (.14 to .29). The differences could be a result of self-reporting differences among black respondents and white respondents, unique conditions affecting black respondents and white respondents, ways in which physicians evaluate both groups, and/or racial differences in the use of medical care. Indeed, white Americans have higher rates of physician visits than do African Americans (National Center for Health Statistics 1993), which suggests that whites should manifest higher morbidity via self reports—precisely what was observed in NHANES-I. Thus, although self-reported morbidity as measured here was equal or superior to physician-evaluated morbidity in a prognostic sense, it is still biased by differences in the use of medical care. Both measures of morbidity may underestimate the true prevalence of disease among African Americans, suggesting that the black/white gap in morbidity may be larger than is commonly assumed.

Kenneth F. Ferraro *is Professor of Sociology and Director of the Gerontology Program at Purdue University. He also serves as Resident Scientist for the National Archive of Computerized Data on Aging at the Inter-university Consortium for Political and Social Research, University of Michigan. His research interests include health status assessment, obesity, and racial differences in health and health service use across the life course. Recent publications appear in* Journal of Gerontology: Social Sciences, Journal of Health and Social Behavior, *and* Social Forces.

Melissa M. Farmer *completed her Ph.D. in sociology at Purdue University in 1998. She is now an assistant researcher with the Department of Health Services and the Jonsson Comprehensive Cancer Center at the University of California, Los Angeles. Her research interests include cancer prevention and early detection, subjective health, disability and functional limitation, stress across the life course, and African American health. Recent publications appear in* Annals of Internal Medicine, Journal of Gerontology: Social Sciences, *and* Journal of Health and Social Behavior.

REFERENCES

Altman, Roy D. 1991. "Classification of Disease: Osteoarthritis." *Seminars in Arthritis and Rheumatism* 20(Supp. 2):40–47.

American Medical Association. 1991. "Hispanic Health in the United States." *Journal of the American Medical Association* 265:248–52.

Anderson, Dallas W., Bruce S. Schoenberg, and Armin F. Haerer. 1988. "Prevalence Surveys of Neurological Disorders: Methodologic Implications of the Copiah County Study." *Journal of Clinical Epidemiology* 41:339–45.

Angel, Ronald and William Gronfein. 1988. "The Use of Subjective Information in Statistical Models." *American Sociological Review* 53: 464–73.

Angel, Ronald and Peter J. Guarnaccia. 1989. "Mind, Body, and Culture: Somatization among Hispanics." *Social Science and Medicine* 28:1229–38.

Blustein, Jan and Beth Weitzman. 1995. "Access to Hospitals with High-Technology Cardiac Services: How Is Race Important?" *American Journal of Public Health* 85:345–51.

Bradburn, Norman M., Lance J. Rips, and Steven K. Shevell. 1987. "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys." *Science* 236:157–61.

Colditz, Graham A., Pauline Martin, Meir J. Stampfer, Walter C. Willett, Laura Sampson, Bernard Rosner, Charles H. Hennekens, and Frank E. Speizer. 1986. "Validation of Questionnaire Information on Risk Factors and Disease Outcomes in a Prospective Cohort of Women." *American Journal of Epidemiology* 123:894–900.

Coughlin, Steven S. 1990. "Recall Bias in Epidemiologic Studies." *Journal of Clinical Epidemiology* 43:87–91.

Ferraro, Kenneth F. and Melissa M. Farmer. 1996. "Double Jeopardy to Health Hypothesis for African Americans: Analysis and Critique." *Journal of Health and Social Behavior* 37:27–43.

Ford, Earl S., Yossi Harel, Gregory Heath, Richard S. Cooper, and Carl J. Caspersen. 1990. "Test Characteristics of Self-Reported Hypertension among the Hispanic Population: Findings from the Hispanic Health and Nutrition Examination Survey." *Journal of Clinical Epidemiology* 43:159–65.

Freidson, Eliot. 1970. *The Profession of Medi-*

cine: A Study of the Sociology of Applied Knowledge. New York: Dodd, Mead.

Gujarati, Damodar N. 1988. *Basic Econometrics*. New York: McGraw-Hill.

Guralnik, J. M., L. P. Fried, J. D. Williamson, and M. C. Hochberg. 1996. "Differences in Ascertainment of Disease Presence by Self-Report and Adjudication." *The Gerontologist* 36(supp. 1):234–35.

Guralnik, Jack M., Linda P. Fried, Eleanor M. Simonsick, Judith D. Kasper, and Mary E. Lafferty, eds. 1995. *The Women's Health and Aging Study*. NIH Pub. No. 95–4009. Bethesda, MD: National Institute on Aging.

Hadaway, C. Kirk, Penny Long Marler, and Mark Chaves. 1993. "What the Polls Don't Show: A Closer Look at U.S. Church Attendance." *American Sociological Review* 58:741–52.

Hall, Judith A., Debra L. Roter, and N. R. Katz. 1988. "Meta-Analysis of Correlates of Provider Behavior in Medical Encounters." *Medical Care* 26:657–75.

Harlow, Sioban D. and Martha S. Linet. 1989. "Agreement between Questionnaire Data and Medical Records: The Evidence for Accuracy of Recall." *American Journal of Epidemiology* 129:233–48.

Heliövaara, M., A. Aromaa, T. Klaukka, P. Knekt, M. Joukamaa, and O. Impivaara. 1993. "Reliability and Validity of Interview Data on Chronic Diseases: The Mini-Finland Health Survey." *Journal of Clinical Epidemiology* 46:181–91.

Himmelstein, David U. and Steffie Woolhandler. 1995. "Care Denied: US Residents Who Are Unable to Obtain Needed Medical Services." *American Journal of Public Health* 85:341–44.

Hummer, Robert A. 1996. "Black-White Differences in Health and Mortality: A Review and Conceptual Model." *Sociological Quarterly* 37:105–25.

Idler, Ellen L. and Ronald J. Angel. 1990. "Self-Rated Health and Mortality in the NHANES-I Epidemiologic Follow-Up Study." *American Journal of Public Health* 80:446–52.

Mutchler, Jan E. and Jeffrey A. Burr. 1991. "Racial Differences in Health and Health Care Service Utilization in Later Life: The Effect of Socioeconomic Status." *Journal of Health and Social Behavior* 32:342–56.

Nadeau, Richard and Richard G. Niemi. 1995. "Educated Guesses: The Process of Answering Factual Knowledge Questions in Surveys." *Public Opinion Quarterly* 59:323–46.

National Center for Health Statistics. 1965. *Health Interview Responses Compared with Medical Records. Vital and Health Statistics*. Series 2, No. 7. USPHS Publication No. 1000. Rockville, MD: National Center for Health Statistics.

———. 1979. "Plan and Operation of the Health and Nutrition Examination Survey: United States—1971–1973." *Vital and Health Statistics*. Series 1, No. 10a. Washington, DC: U.S. Government Printing Office.

———. 1993. *Health, United States, 1992*. Washington, DC: U.S. Government Printing Office.

National Center for Health Statistics, and Bruce B. Cohen, Helen E. Barbano, Christine S. Cox., Jacob J. Feldman, Fanchon F. Finucane, Joel C. Kleinman, and Jennifer H. Madans. 1987. "Plan and Operation of the NHANES I Epidemiologic Followup Study, 1982–1984." *Vital and Health Statistics*. Series 1, No. 22. Washington, DC: U.S. Government Printing Office.

Pecoraro, Roger E., Thomas S. Inui, Wei S. Chen, Diana K. Plorde, and Jacob L. Heller. 1979. "Validity and Reliability of a Self-Administered Health History Questionnaire." *Public Health Reports* 94:231–38.

Presser, Stanley and Michael Traugott. 1992. "Little White Lies and Social Science Models: Correlated Response Errors in a Panel Study of Voting." *Public Opinion Quarterly* 56:77–86.

Shah, Babubhai V., Beth G. Barnwell, and Gayle S. Bieler. 1997. *SUDAAN User's Manual*. Rel. 7.5. Research Triangle Park, NC: Research Triangle Institute.

Smith, Kevin W., Sarah A. McGraw, Sybil L. Crawford, and John B. McKinlay. 1993. "Do Blacks and Whites Differ in Reporting Rose Questionnaire Angina? Results from the Boston Health Care Project." *Ethnicity and Disease* 3:278–89.

Wolinsky, Fredric D., Benigno E. Aguirre, Lihjiuan Fann, Verna M. Keith, Connie L. Arnold, John C. Niederhauer, and Kathy Dietrich. 1989. "Ethnic Differences in the Demand for Physician and Hospital Utilization among Older Adults in Major American Cities: Conspicuous Evidence of Considerable Inequalities." *Milbank Quarterly* 67:412–49.

World Health Organization. 1967. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death: Based on Recommendations of the Eighth Revision Conference, 1965, and Adopted by the Nineteenth World Health Assembly*. Geneva, Switzerland: World Health Organization.