Physician-Evaluated and Self-Reported Morbidity for Predicting Disability

Kenneth F. Ferraro, PhD, and Ya-ping Su, PhD

ABSTRACT

Objectives. This study compared the predictive validity of physician-evaluated morbidity and self-reported morbidity on disability among adults.

Methods. Subjects from a large national survey (n = 6913) received a detailed medical examination by a physician and were asked about the presence of 36 health conditions at baseline. Disability measured 10 and 15 years later was regressed on the morbidity measures and covariates with tobit models.

Results. Although physician-evaluated morbidity and self-reported morbidity were associated with greater disability, self-reports of chronic nonserious illnesses manifested greater predictive validity. Disability was also higher for obese subjects and those of lower socioeconomic status.

Conclusions. The findings demonstrate the predictive utility of self-reported morbidity measures on functional disability. (*Am J Public Health*. 2000;90:103–108)

The presumption among most epidemiologists and health researchers is that data from physicians are superior to those collected from subjects. Medical examination data are often seen as the "gold standard," and subject reports that are not corroborated by data from physician examinations are typically regarded as "inaccuracies."[1] If survey research subjects report disease not detected by a physician examination, the subjects are typically seen as "somatizing"[2] or pessimistic in evaluating their health.[3] In contrast, when a physician identifies a disease that a subject does not report, the incongruity usually is seen as resulting from the subject's optimistic outlook, a lack of access to medical care, or denial of illness. These may be reasonable assumptions under some circumstances, but the science of self-report suggests that the relationships are probably more complex.[4–6] Except for autopsy, physicians rely—at least in part—on subject information when formulating a diagnosis. Subject reports of disease, in turn, often hinge on recall of what physicians communicated in medical encounters.

Although it is reasonable to regard data from physicians as the gold standard for health measurement, this research compared morbidity reports from survey respondents with medical evaluations of morbidity by physicians in a national health survey. A number of studies comparing data from these 2 sources have focused on detection of bias or recall error for self-reported morbidity.[7] This study posed a different question: Which data source is more useful in a prognostic sense? The overarching research question is whether *physician-evaluated morbidity* and *self-reported morbidity* manifest predictive validity in models of functional disability over time. If both are predictive, which one has greater predictive validity?

Functional disability is an important outcome to consider because of its pivotal role in shaping health trajectories,[8,9] health care needs and costs,[10–12] and mortality risk.[13] Although we are not aware of any study that compared the influence of physician-evaluated and self-reported morbidity on disability, a growing body of research is showing the utility of self-reported data for predicting disability.[13,14]

Morbidity and Disability

A substantial and growing body of research shows clearly that morbidity is among the most important predictors of most measures of functional impairment and disability. The literature shows consistently that morbidity is a major determinant of active life expectancy[15] and the disablement process.[16,17] Moreover, morbidity is fundamental to most models of the structure of health status and health-related quality of life.[18,19] Of course, the strength of the association between morbidity and disability is dependent on valid and reliable measurement of each.

Disability is typically measured with a battery of self-report questions about physical function (e.g., activities of daily living). These questions are usually specific to the performance of functional tasks, some of which could be readily assessed during a face-to-face interview. Morbidity also is measured most often in survey research by self-report. A question about a disease may appear to be a simple query, but the response can be fairly complex. With regard to the true prevalence of morbidity—unlike disability or health

Kenneth F. Ferraro is with the Department of Sociology, Purdue University, West Lafayette, Ind. Ya-ping Su is with the Lighthouse International, New York, NY.

Correspondence and requests for reprints should be sent to Kenneth F. Ferraro, PhD, Department of Sociology, 1365 Stone Hall, Purdue University, West Lafayette, IN 47907-1365 (e-mail: ferraro@purdue.edu).

This article was accepted July 9, 1999.

assessments—knowledge about a condition is usually conferred from medical examinations.[20] Subjects may suspect that they have a disease but may not have received a definitive diagnosis. Thus, a report of a disease in such a case could be seen as either a report of preclinical morbidity or somatizing.[2]

Even more basic to the measurement of morbidity is the implied use of medical services. Although disability indicators such as carrying a 10-pound bag of groceries or picking up a penny rely only on subject report, morbidity reports are, to some degree, dependent on interaction with the medical care system. Contact with physicians either validates or confutes a subject's presumption. Important economic and cultural selection processes also shape contact with physicians.[21–23] In short, subject reports of morbidity may be biased by differences in subjects' knowledge of diseases, health service use, and health communication.[24,25] If morbidity is systematically underestimated or overestimated in health surveys, then the link between morbidity and disability as well as other dimensions of health-related quality of life also may have been misrepresented in previous studies.

Dozens of studies have compared morbidity measures from subjects and physicians,[1,26–29] but little research has compared the 2 sources of morbidity data in a prognostic sense. If physician-evaluated morbidity is the best available measure of morbidity in social surveys, then it is reasonable to expect it to be predictive of health-related quality of life at later points in time. One recent study[30] compared the predictive validity of 2 types of morbidity on self-assessed health and survival and found that subject-reported morbidity was not inferior to physician-evaluated morbidity. The present study adds to this line of research by considering disability as the outcome. This research seeks to contribute to the understanding of the accuracy of morbidity information gathered from survey respondents as well as the relationship between morbidity and disability in models of health-related quality of life.

Methods

Sample

Data from the National Health and Nutrition Examination Survey I (NHANES I) were used in this research.[31] NHANES is unique because of its collection of both medical examination and survey interview data. For the present research, a longitudinal component of the study allowed an analysis of the predictive validity of the 2 types of morbidity on disability over a 15-year observation period. In addition to providing detailed health information, NHANES is an important source of information on health behaviors and status characteristics likely to be important to disability. One limitation of the archive is that the medical examination data were collected only during the baseline interview. Therefore, change in the 2 types of morbidity cannot be examined. The strategy implemented here was to examine the predictive validity of the 2 types of morbidity on disability over the 15-year study period.

The baseline NHANES I was conducted from 1971 to 1975. The sampling design involved a multistage, stratified probability sample of noninstitutionalized persons aged 25 to 74 years.[31] This study used data from the baseline survey and the Epidemiologic Follow-Up Study completed during 1982 through 1984 (wave 2) and 1987 (wave 3).[32,33]

The analyses were completed on the NHANES I subsample that was administered the "detailed component," including the extensive medical examination and the Health Care Needs Questionnaire, at baseline (n = 6833). The sample used in this study was composed of 5955 White (87.2%) and 878 Black respondents (12.8%) at baseline; the unweighted data were used throughout, consistent with the recommendations of others.[34,35] The percentage of patients receiving the detailed component at baseline and traced through follow-up was very high (92.6% of the survivors at wave 2 and 96.5% of the survivors at wave 3). The number of cases lost to death, tracing, and refusal to participate was 1644 by wave 2 and 2078 by wave 3 (approximately 15 years later). As is discussed later in this article, selection bias models are used to account for attrition in estimates of disability at the follow-up surveys.[36] Given the large number of cases available for analyses, a more conservative probability level of .01 was selected for statistical significance.

Measurement

Each subject who received the detailed component of the survey was given an extensive medical examination. NHANES representatives visited the households of sample members to administer a basic health survey and a medical history interview and to arrange for the subject to come to a nearby mobile examination center (i.e., 3 connected trailers). The NHANES staff trained physicians in the conduct of the examination, and physicians reviewed the medical history questionnaire for each subject the day before the scheduled examination.[31] Physicians also had ready access to the results of numerous NHANES laboratory tests (e.g., hematology, goniometry, and various x-rays). Thus, physicians were given extensive information on the subject before conducting the medical examination.

Physician-evaluated morbidity was assessed in 2 ways. First, findings from the full examination—drawing from both the laboratory tests and the physician examination—were categorized by the physician according to the *International Classification of Diseases, Eighth Revision* (ICD-8).[37] Following the procedures of others,[38] physicians used 15 ICD morbidity categories to summarize findings: (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. The ICD codes are widely used and were designed to be truly comprehensive in reliably classifying many types of disorders.[37]

Second, findings from the general medical examination were coded into 8 categories by the physicians. These categories are (1) head, eyes, ears, nose, and throat; (2) thyroid; (3) chest; (4) cardiovascular; (5) abdominal; (6) musculoskeletal; (7) neurological; and (8) skin. The general medical examination (8 domains) was designed to uncover findings based on anatomy and function. The 2 protocols represent different classification schemes, but each measures physician-evaluated morbidity. Findings from each measure were first binary coded (0, 1) and then summed; the simple correlation between them was 0.42.

NHANES staff collected most of the remaining health measures used in the analysis in one of the interviews with the subject. Self-reported morbidity was derived from a checklist-type question designed to identify which illnesses respondents had. Respondents were asked the following question: "Has a doctor ever told you that you have [hypertension or high blood pressure, etc]?" (Thirty-six conditions were presented to all subjects.) The answer to such a question is not a report about how one feels about a specific condition but a report of a condition based on a medical examination. Unlike some surveys that ask whether a person has a condition, NHANES hinged the question on evaluation by a physician.

Each condition was coded as a binary variable (0, 1). The conditions were then classified into those that were life threatening or serious and all remaining conditions.[39] Serious conditions included cancer, diabetes, heart failure (attack or trouble), hypertension, and stroke. Chronic nonserious conditions included arthritis, asthma, bone fracture, cataracts, gout, psoriasis, and ulcer. The serious and chronic nonserious conditions were then summed separately; the simple correlation between them was 0.24. It is clear that differences in the protocol and question wording for measuring physician-evaluated and self-reported morbidity preclude disease-by-disease comparisons for all conditions. (Respondents were asked about only 36 conditions, but physicians had hundreds of ICD codes from which to choose. Thus, beyond differences in question wording, aggregation bias is possible when trying to make disease-by-disease comparisons.) Despite this limitation of NHANES I, the respective procedures for physicians and subjects are among the most widely used to measure morbidity in health surveys. The correlations between the 2 physician-evaluated and 2 self-reported morbidity measures ranged from 0.22 to 0.26.

The measures of functional disability were not identical from baseline to follow-up surveys. In fact, no measure of disability was collected during the baseline interview. Subjects were, however, asked whether they normally experienced pain in the joints, back or neck, and hip or knees. Respondents were told to identify chronic pain—defined as pain that lasted for at least 1 month—that occurred either while they were at rest or while moving the respective joints. The interviewer solicited responses for each of the 3 skeletal areas (coded as 1 = yes, 0 = no). The α coefficient of reliability for the additive index of the 3 items was .74. The index is a fairly reliable measure of chronic pain but was used here as a control variable in predicting disability, not a measure of disability. Change in disability from the baseline survey cannot be determined.

For the second and third waves, the Stanford Health Assessment Questionnaire Disability Index was used. It asks very specific questions; for example, "Please tell me if you have no difficulty, some difficulty, much difficulty or are unable to do these activities at all when you are by yourself and without the use of aids [e.g., lift and carry a full bag of groceries]." Responses for each item ranged from 1 (no difficulty) to 4 (unable to do). The original index included 26 items, but a few items were either deleted or modified in the 1987 interview.[40,41] A total of 21 items (common to both waves) were used for the disability index. Thus, the index ranged from 21 to 84 at both waves but had a mean slightly greater than 23.

At least 50% of the subjects at each wave scored 21 (no disability), whereas fewer than 1% of the respondents scored 84 (unable to do any tasks). The items comprised a wide range of functions, including dressing and grooming, hygiene, eating, walking, reaching, gripping, and activities (errands). The α coefficient of reliability for the 21 items was .96 in wave 2 and .98 in wave 3. Missing data on 6 or fewer items of the 21-item index were recoded to group means defined by age, sex, and race grouping. If respondents missed more than 6 items, the respondents were treated as missing on the index.[42,43]

The remaining independent variables span a broad range of factors related to morbidity and disability, either directly or indirectly. These include indicators of health risk behaviors such as obesity and smoking. Obese subjects were identified by the physician during the medical examination. Smoking was based on self-report of consumption of cigarettes, cigars, and pipe tobacco at the time of the interview and during one's lifetime. Variables related to social class included education, availability of private health insurance, Medicaid status, and family income. The measurement of the remaining independent variables was fairly straightforward and is described in Table 1 (binary variables were coded 0 and 1).

Analytic Plan

As noted earlier, the disability index had a large percentage of persons in this sample with no disability (51% of wave 2 and 50% of wave 3 subjects had no disability). Given the skewed distribution of the disability measures, tobit regression models were estimated. Although case tracing and reinterview rates were high in the NHANES I follow-ups, attrition in longitudinal analyses may influence sample estimates of predictor variables and lead to bias in the estimates of the true relationship between the independent variables and disability. Thus, selection bias models, originally developed by Heckman, were used to correct parameter estimates for differential selectivity due to death, refusal to participate, or inability to trace.[44–46] The procedure is to (1) estimate a probit model to distinguish subjects who participate from those who do not, and (2) use the probit results to create a selection (hazard) instrument based on the inverse Mills ratio, adding the selection instrument to the regression model of interest (as an independent variable). The analyses were completed in LIMDEP, which provides for the use of tobit models while accounting for the selection process.[47]

Results

The probit model estimating attrition between the first two survey waves showed that subjects most likely to drop out of the analysis were Black, older, and male and had less education than other subjects. (The selection equation also included missing on income at wave 1, a variable not included in the substantive equation, but its effect was nonsignificant.)

The results of the tobit analysis predicting disability at wave 2 are shown in the first column of Table 1 and include the term for the selection effect (λ). The analysis shows that the relationships between the independent variables and disability were not significantly affected by the adverse selection during the first 10 years of the study. Physician-evaluated morbidity and self-reported morbidity were significantly related to disability at wave 2, and all 4 measures manifested positive relationships. The magnitude of the effects was greater for self-reported morbidity, as reflected in tests of the differences in slopes (*t* test values, not presented, ranged from –3.82 to –8.21).[48] No differences in disability were found between White and Black subjects at wave 2, but disability was higher for women than for men. Disability was also higher for those with less income, persons on Medicaid, obese subjects, and smokers.

The second column of Table 1 presents the results for disability by wave 3 (about 15 years after the first interview). The probit model estimating attrition between wave 1 and wave 3 found that subjects more likely to drop out of the analysis were older, Black, and male; had less income than other subjects; and did not report income at wave 1. The results for wave 3 disability show even more clearly the superiority of self-reported morbidity to physician-evaluated morbidity. ICD morbidity is not significant for predicting disability by wave 3, but each of the remaining morbidity measures manifests a positive relationship. As in the previous analysis, tests of slope differences showed that measures of self-reported morbidity were stronger predictors (*t* test values ranged from –2.73 to –5.11).[48] A test of slope differences between serious illness and chronic nonserious illness showed that the latter had a significantly stronger effect on disability. The sex difference in disability is no longer manifest at wave 3. Disability was higher, however, for older people, those with limited education and income, obese subjects, and smokers.

The final equation, shown in the third column of Table 1, added disability at wave 2 so that change in disability between wave 2 and wave 3 could be predicted. Again, the superiority of the self-reported morbidity measures is clear, and, as one would expect, chronic illness is more consequential to change in functional disability than serious or life-threatening morbidity. Disability was more likely to increase among older persons,

TABLE 1—Predicting Disability Among Adults at Waves 2 and 3[a] of the National Health and Nutrition Examination Survey I: Epidemiologic Follow-Up Study (Tobit Regression Estimates)

Variable (Coding)	Disability, Wave 2 (n = 5121) b (SE)	Disability, Wave 3 (n = 4342) b (SE)	Disability, Wave 3 (n = 4342) b (SE)
Self-reported morbidity			
Serious illness (0–4; 4 = 4+)	2.32 (0.39)***	3.12 (0.64)***	2.16 (0.59)**
Chronic nonserious illness (0–10; 10 = 10+)	4.39 (0.49)***	5.47 (0.78)***	3.19 (0.73)***
Physician-evaluated morbidity			
ICD findings (0–5; 5 = 5+)	0.90 (0.24)**	0.96 (0.38)	0.70 (0.35)
General medical examination (0–5; 5 = 5+)	0.82 (0.23)**	1.39 (0.37)**	1.01 (0.35)*
Black (1 = Black)	–2.11 (3.31)	5.86 (2.41)	6.33 (2.22)*
Self-reported pain (0–3; α = .74)	0.13 (0.23)	–0.48 (0.38)	–0.63 (0.35)
Age, y (24–77)	0.18 (0.09)	0.46 (0.12)**	0.38 (0.12)**
Female (1 = female)	3.71 (1.30)*	2.00 (1.68)	0.84 (1.55)
Lives alone (1 = lives alone)	0.15 (0.86)	1.30 (1.42)	1.49 (1.31)
Widow (1 = widowed)	0.61 (0.95)	1.93 (1.57)	1.32 (1.45)
Community type (1 = rural)	–0.82 (0.51)	0.07 (0.81)	0.68 (0.75)
Education (0–7; 7 = graduate school)	–0.30 (0.67)	–2.57 (0.70)**	–2.33 (0.65)**
Income, $ (1–12; 12 ≥ 25 000)	–0.58 (0.11)***	–0.65 (0.18)**	–0.40 (0.17)
Private insurance (1 = yes)	0.31 (0.71)	–1.33 (1.12)	–1.79 (1.03)
Medicaid card (1 = yes)	3.67 (1.27)**	–1.59 (2.27)	–3.96 (2.13)
Regular physician (1 = yes)	0.39 (0.73)	–0.34 (1.18)	–0.39 (1.09)
Obese (1 = yes)	1.84 (0.57)**	4.31 (0.90)***	2.78 (0.84)**
Smoker (1 = current smoker)	1.93 (0.52)**	2.42 (0.85)*	1.94 (0.79)
Past smoker (1 = yes)	0.78 (0.61)	0.41 (0.99)	0.54 (0.91)
Disability, wave 2			1.05 (0.06)***
Selection (λ)	10.49 (17.25)	–14.36 (10.68)	–14.00 (9.84)
Intercept	–22.84 (5.56)***	–21.78 (3.87)***	–19.10 (3.57)***
Log likelihood	–7895.38	–5975.09	–5798.38

Note. ICD = *International Classification of Diseases, Eighth Revision.*
[a]The first 2 equations examine disability at the respective waves. The final equation specifies wave 2 disability as an independent variable to examine change in disability by wave 3.
**P* < .01; ***P* < .001; ****P* < .0001.

Black adults, those with limited education, obese subjects, and smokers.

The final stage of the analysis involved examining possible interactions by race. It was anticipated that physician-evaluated morbidity would be more predictive among White than Black subjects. Separate sample analyses and tests of interaction terms did not support this expectation. None of the interaction terms between Black subjects and the 4 morbidity measures were significant in parallel tobit models.

Discussion

Epidemiologists and health behavior scientists often rely on self-reported measures of morbidity for studies of the sequelae of disease, including disability, health perceptions, and health service use.[19] The purpose of this research was to compare the predictive validity of self-reported morbidity measures with that of those collected by physicians in a major health survey. The results showed that measures of both physician-evaluated morbidity and self-reported morbidity were significant predictors of disability 10 years after the baseline survey but that the strongest effects were due to the self-reported measures, especially chronic nonserious illness. Moreover, for the 15-year follow-up, physician-evaluated morbidity based on the ICD was nonsignificant. Thus, the expectation that physician-evaluated morbidity measures would be superior to self-reported morbidity for predicting disability was not supported. In fact, even when both types of morbidity were significant, the magnitude of the effect was greater for self-reported morbidity.

Although physician-evaluated data are typically considered the gold standard for measuring morbidity, this assumption may merit reconsideration when medical examination data are drawn from health surveys. This is the second report from the NHANES I Epidemiologic Follow-Up Study showing that measures of self-reported morbidity are not inferior to physician-evaluated data in a prognostic sense.[30] In addition, evidence that physicians underestimate or fail to recognize disabilities reported by patients suggests that relying solely on physician-evaluated data may lead to biased predictions of health trajectories for both survey respondents and patients.[49] Nevertheless, there have been important changes in both the procedures for physical examinations in health surveys and the practice guidelines since the baseline of NHANES I.[50,51] Thus, whether these findings can be replicated with more recent data remains to be seen.

With disability as the outcome of consideration here, evidence of the value of self-reported morbidity in comparison with physician evaluations in survey research is growing. Physician-evaluated morbidity from clinical records may be an entirely different matter. It may be superior to self-reported morbidity because of the more extensive and enduring communication between physician and patient.[52,53] Yet, accessing medical records and ensuring standardization across clinical settings are frequent barriers to using those data.[29] Moreover, clinical data do not represent populations, only those who consume services. For many issues, therefore, especially those related to health inequality, sample surveys provide the most appropriate data. Although basic (onetime) medical evaluations in epidemiologic surveys may underestimate morbidity, the procedures used in NHANES were fairly comprehensive.

The fact that self-reported morbidity is a stronger predictor of disability than is physician-evaluated morbidity is even more intriguing given the laboratory and medical history information provided to physicians before the medical examination. The stronger association between self-reported morbidity and disability may be partly caused by an underlying tendency to complain about both disease and disability. Because no systematic way to test such a thesis exists in NHANES I, such a mechanism should be considered in subsequent research.

Despite the limitations of the data, these findings bolster confidence in the use of self-report measures in the health sciences and in models of the relationship between morbidity and disability. It should be recalled, however, that the condition checklist in the NHANES I was fairly extensive. Whether the shorter lists, widely used in other national surveys, are equally useful is unclear. Whereas the shorter lists emphasize the serious illnesses, it would appear that they should also manifest considerable predictive validity for survival but perhaps not for disability and other dimensions of health-related quality of life.

The indicators of self-reported disease in the present analysis were grouped into serious illnesses such as cancer and heart trouble and chronic nonserious conditions such as arthritis and back trouble. This distinction proved useful in modeling disability and change in disability over time. Although analyses of these data for mortality show the importance of serious illness,[30] the present analysis focused on disability and found that chronic nonserious illness was the strongest of the morbidity predictors. The grouping of illnesses as done here merits further examination, but this analysis confirms the utility of this approach with self-reported morbidity. □

Contributors

K. F. Ferraro planned the study, supervised the data analysis, and wrote the paper. Y. Su performed all statistical analyses and contributed to the writing of the paper. Both authors contributed to the identification and operationalization of the variables, development of the analytic methods, and interpretation of the data.

Acknowledgments

This research was supported by a grant from the National Institute on Aging (AG 11705).

We appreciate the assistance of Ronald Angel and Ellen Idler with the construction of selected morbidity variables and Melissa Farmer and Priya Rajagopalan with computing.

The data used in this paper were made available by the Inter-University Consortium for Political and Social Research. Neither the collector of the original data nor the Consortium bears any responsibility for the analyses or interpretations presented here.

References

1. Kriegsman DMW, Penninx WJH, van Eijk JTM, Boeke AJP, Deeg DJH. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly: a study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J Clin Epidemiol*. 1996;49:1407–1417.
2. Angel R, Guarnaccia PJ. Mind, body, and culture: somatization among Hispanics. *Soc Sci Med*. 1989;28:1229–1238.
3. Maddox GL. Some correlates of differences in self-assessment of health status among the elderly. *J Gerontol*. 1962;17:180–185.
4. Markides KS, Lee DJ, Ray LA, Black SA. Physicians' ratings of health in middle and old age: a cautionary note. *J Gerontol B Psychol Sci Soc Sci*. 1993;48(suppl):S24–S27.
5. Nadeau R, Niemi RG. Educated guesses: the process of answering factual knowledge questions in surveys. *Public Opin Q*. 1995;59: 323–346.
6. Presser S, Traugott M. Little white lies and social science models: correlated response errors in a panel study of voting. *Public Opin Q*. 1992;56:77–86.
7. Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol*. 1990;43:87–91.
8. Ferraro KF, Farmer MM, Wybraniec JA. Health trajectories: long-term dynamics among black and white adults. *J Health Soc Behav*. 1997;38: 38–54.
9. Palmore EB, Burchett BM. Predictors of disability in the final year of life. *J Aging Health*. 1997;9:283–297.
10. Lubitz J, Prihoda R. The use and costs of Medicare services in the last two years of life. *Health Care Fin Rev*. 1984;5:117–131.
11. Salive ME, Guralnik JM. Disability outcomes of chronic disease and their implications for public health. In: Hickey T, Speers MA, Prohaska TR, eds. *Public Health and Aging*. Baltimore, Md: Johns Hopkins University Press; 1997:87–106.
12. Van Den Bos G. The burden of chronic diseases in terms of disability, use of health care and health life expectancies. *Eur J Public Health*. 1995;5:29–34.
13. Greiner PA, Snowdon DA, Greiner LH. The relationship of self-rated function and self-rated health to concurrent functional ability, functional decline, and mortality: findings from the nun study. *J Gerontol B Psychol Sci Soc Sci*. 1996;51B:S234–S241.
14. Idler EL, Kasl SV. Self-ratings of health: do they also predict change in functional ability? *J Gerontol B Psychol Sci Soc Sci.* 1995;50B: S344–S353.
15. Crimmins EM, Hayward MD, Saito Y. Differentials in active life expectancy in the older population of the United States. *J Gerontol B Psychol Sci Soc Sci*. 1996;51B:S111–S120.
16. Lawrence R, Jette A. Disentangling the disablement process. *J Gerontol B Psychol Sci Soc Sci.* 1996;51B:S173–S182.
17. Verbrugge LM, Jette AM. The disablement process. *Soc Sci Med*. 1994;38:1–14.
18. Johnson RJ, Wolinsky FD. The structure of health status among older adults: disease, disability, functional limitation, and perceived health. *J Health Soc Behav*. 1993;34:105–121.
19. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. *JAMA*. 1995;273:59–65.
20. Brown P. Naming and framing: the social construction of diagnosis and illness. *J Health Soc Behav*. 1995;Spec No:34–52.
21. Himmelstein DU, Woolhandler S. Care denied: US residents who are unable to obtain needed medical services. *Am J Public Health.* 1995;85: 341–344.
22. Lieu TA, Newacheck PW, McManus MA. Race, ethnicity, and access to ambulatory care among US adolescents. *Am J Public Health*. 1993;83: 960–965.
23. Wolinsky FD, Aguirre BE, Fann L, et al. Ethnic differences in the demand for physician and hospital utilization among older adults in major American cities: conspicuous evidence of considerable inequalities. *Milbank Q*. 1989;67: 412–449.
24. Hall JA, Roter DL, Katz NR. Meta-analysis of correlates of provider behavior in medical encounters. *Med Care*. 1988;26:657–675.
25. Roter DL, Hall JA. *Doctors Talking With Patients/Patients Talking With Doctors*. Westport, Conn: Auburn House; 1992.
26. Berkanovic E, Hurwicz M, Lachenbruch PA. Concordant and discrepant views of patients' physical functioning. *Arthritis Care Res*. 1995; 8:94–101.
27. Guralnik JM, Fried LP, Simonsick EM, Kasper JD, Lafferty ME, eds. *The Women's Health and Aging Study: Health and Social Characteristics of Older Women With Disability*. Bethesda, Md: National Institute on Aging; 1995. NIH publication 95-4009.
28. Heliövaara M, Aromaa A, Klaukka T, Knekt P, Joukamaa M, Impivaara O. Reliability and validity of interview data on chronic diseases: the Mini-Finland Health Survey. *J Clin Epidemiol*. 1993;46:181–191.
29. Harlow SD, Linet MS. Agreement between questionnaire data and medical records: the evidence for accuracy of recall. *Am J Epidemiol*. 1989;129:233–248.
30. Ferraro KF, Farmer MM. Utility of health data from social surveys: is there a gold standard for measuring morbidity? *Am Sociol Rev.* 1999; 64:303–315.
31. National Center for Health Statistics. Plan and operation of the Health and Nutrition Examination Survey: United States—1971–1973. *Vital Health Stat 1*. 1973;10a. DHEW publication PHS 79-1310.
32. National Center for Health Statistics, Cohen BB, Barbano HE, Cox CS, et al. Plan and operation of the NHANES I epidemiologic followup study, 1982–1984. *Vital Health Stat 1*. Washington, DC: Government Printing Office; 1987; No. 22.
33. Barbano HE, Cohen BB, Kleinman JC, Madans JH, Feldman JJ, Cornoni-Huntley JC. Objective, design, and implementation. In: Cornoni-Huntley JC, Huntley RR, Feldman JJ, eds. *Health Status and Well-Being of the Elderly: National Health and Nutrition Examination Survey–I Epidemiologic Follow-Up Study*. New York, NY: Oxford University Press; 1990:3–17.
34. Feldman JJ, Makuc DM, Kleinman JC, Cornoni-Huntley J. National trends in educational differentials in mortality. *Am J Epidemiol.* 1989; 129:919–933.

35. Winship C, Radbill L. Sampling weights and regression analysis. *Sociol Methods.* 1994;23: 230–257.
36. Stolzenberg RM, Relles DA. Tools for intuition about sample selection bias and correction. *Am Sociol Rev*. 1997;62:494–507.
37. World Health Organization. *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; 1967.
38. Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I Epidemiologic Follow-Up Study. *Am J Public Health.* 1990;80:446–452.
39. Ferraro KF, Farmer MM. Double jeopardy, aging as leveler, or persistent health inequality: a longitudinal analysis of white and black Americans. *J Gerontol B Psychol Sci Soc Sci*. 1996;51B:S319–S328.
40. Foley DJ, Branch LG, Madans JH, Brock DB, Guralnik JM, Williams TF. Physical function. In: Cornoni-Huntley JC, Huntley RR, Feldman JJ, eds. *Health Status and Well-Being of the Elderly: National Health and Nutrition Examination Survey–I Epidemiologic Follow-Up Study.* New York, NY: Oxford University Press; 1990:221–236.
41. Thompson PW, Pegley FE. A comparison of disability measured by the Stanford Health Assessment Questionnaire (HAQ) Disability Scales in male and female rheumatoid outpatients. *Br J Rheumatol*. 1991;30:298–300.
42. Rubin DA. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: John Wiley & Sons Inc; 1987.
43. Little RJ, Schenker N. Missing data. In: Arminger G, Clogg CC, Sobel ME, eds. *Handbook of Statistical Modeling for the Social and Behavioral Sciences.* New York, NY: Plenum Press; 1995:39–75.
44. Berk R. An introduction to sample selection bias in sociological data. *Am Sociol Rev*. 1983;48:386–398.
45. Winship C, Mare RD. Models for sample selection bias. *Annu Rev Sociol*. 1992;18:327–350.
46. Heckman JJ. Sample selection bias as a specification error. *Econometric.* 1979;47:153–161.
47. Greene WH. *LIMDEP Version 7.0: User's Manual*. Bellport, NY: Econometric Software; 1995.
48. Gujarati DN. *Basic Econometrics.* New York, NY: McGraw-Hill; 1988.
49. Calkins DR, Rubenstein LV, Cleary PD, et al. Failure of physicians to recognize functional disability in ambulatory patients. *Ann Intern Med*. 1991;114:451–454.
50. Joint National Committee. The sixth report of the Joint National Committee on Detection, Evaluation and Treatment of High Blood Pressure. *Arch Intern Med*. 1997;157:2413–2445.
51. National Heart, Lung and Blood Institute. *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults.* Bethesda, Md: National Institutes of Health; 1998.
52. Atkinson P. *Medical Talk and Medical Work.* London, England: Sage Publications; 1995.
53. Waitzkin H. *The Politics of Medical Encounters.* New Haven, Conn: Yale University Press; 1991.