Issues in Survey Data on Medical Practice: Some Empirical Comparisons

HOWARD P. GREENWALD, PhD
L. GARY HART, PhD

Dr. Greenwald is Associate Professor and Academic Coordinator, Graduate Programs in Health Services Administration, School of Public Administration, Sacramento Public Affairs Center, University of Southern California. Dr. Hart is a Research Associate, Department of Family Medicine, University of Washington.

Tearsheet requests to Howard P. Greenwald, PhD, Health Services Administration Program, VKC-263, School of Public Administration, University of Southern California, University Park, Los Angeles, CA 90089-0041.

Synopsis

In recent years, researchers and policymakers have used data from large-scale surveys of physicians to address important issues. A review of several of these surveys explores potential problems in this method of gathering data on physicians' services. To obtain a better grasp of the limitations such problems may pose, we examine several recent surveys, comparing response rates and survey findings, and in one survey the reliability of individual items.

Response rates appear highly sensitive to differences in the approaches made to respondents and their perceptions of the goals of individual investigations. Reliability of survey items seems to depend on the specificity of information requested. Variation in the findings from different surveys may occur for many reasons, but is most likely to be found in response to items whose presentation differs in each survey's research instrument.

Data from these surveys appear clearly useful for some important purposes. The large-scale medical practice survey seems particularly valuable in generating an understanding of differences among specialties in resources used in the delivery of care. Nevertheless, researchers and policymakers must understand the steps necessary to obtain reliable results and possible limitations in the accuracy of findings to make the best use of survey methodology as applied to medical practice.

THE PAST DECADE HAS WITNESSED a growing concern about medical practice in the United States. Related issues have motivated numerous research efforts, some of the most significant taking the form of large-scale surveys of physicians' activities. Surveys of this kind have formed the basis of studies of physician productivity (*1,2*), resources used by physicians in different specialties (*3-5*), and the role of individual specialties in providing ambulatory care (*6*). These surveys have contributed to the deliberations of key planning bodies, such as the Graduate Medical Education National Advisory Committee (*7*). By examining several important studies, we indicate the problems that may occur in the large-scale survey and specify issues that investigators should consider in using available data or planning future survey efforts.

Beginning in the 1950s, systematic surveys of physicians' activities have steadily increased in size and scope. Peterson (*8*), Clute (*9*), and their colleagues directly observed physician-patient encounters to evaluate the quality of patient care. Marsland (*10*) and coworkers recruited physicians to report directly on the care that they provided to their patients. Confined to restricted localities, these studies each covered about 100 physicians and several thousand physician-patient encounters. Those conducting the "second generation" of medical practice surveys selected much larger samples of physicians from national registries and relied on self-reports by respondents to obtain data on practice characteristics and patient encounters. Examples include the National Center for Health Statistics' National Ambulatory Care Survey (NAMCS), the University of Southern California (USC) Medical Activities and Manpower Projects' physician surveys, the Medical Economics Continuing Survey (MEDECON), the American College of Surgeons and American Surgical Association's Study of Surgical Services for the United States, and the American Medical Association's Periodic Survey of Physicians (PSP).

Basic questions about methods may be asked concerning all these studies. Clute raises the issue of reliability, suggesting that two observers, even using the same explicit standards of care, could rate a given practitioner differently. Marsland and

Comparison of three recent surveys of medical practice

Characteristic	NAMCS	USC	MEDECON
Sample source	AMA Masterfile, AOA Masterfile	AMA masterfile	AMA Masterfile
Specialties	All except radiology, anesthesiology, and pathology.	24 direct patient care (medical and surgical).	All
Physician types	Office based, patient care, non-Federal.	All non-Federal except first year medical residents.	Office-based, non-Federal
Geographic extent	Contiguous United States	United States and Puerto Rico	All United States
Smallest geographic detail for analysis.	4 U.S. regions	State (county with permission).	4 U.S. regions
When conducted	Yearly	1976–78	Yearly
Period of observation	52 weeks	Selected weeks	Unknown
Number of respondents	2,000 in 1977	10,000	6,000
Number of patient encounters	51,000 (yearly)	368,000	None
Instruments	Practice questionnaire, log diary.	Practice questionnaire, log diary, log of telephone encounters, weekly activity overview.	Practice questionnaire

NOTE: NAMCS National Center for Health Statistics' National Ambulatory Care Survey; USC University of Southern California Medical Activities and Manpower Projects' physician surveys; MEDECON Medical Economics Continuing Survey; AMA American Medical Association; AOA American Osteopathic Association.

colleagues' reliance on physician self-reports raises similar questions, leaving open the possibility that individual physicians might introduce unknown bias (for example, different diagnoses of similar biological conditions) in reporting on their own activities. In addition to these difficulties, later surveys may encounter serious problems in response rate and unrecognized bias as investigators seek to study ever larger populations and have less direct contact with subjects.

In this article we concentrate on issues, potential problems, and possible sources of inaccuracy in the more recent investigations. First, we compared response rates obtained by NAMCS, USC, and MEDECON. The response rate is a basic indicator of the feasibility of the large-scale medical practice survey. Second, we examined the reliability of items included in the USC investigation, the only survey for which reliability data were available to the authors. In this context, reliability refers to the tendency of individual survey items to elicit the same responses when administered to the same subjects on different occasions. Third, we compared selected findings from NAMCS and USC to indicate the degree to which each replicates the other's results. Both reliability of individual survey items and replication of findings by separate surveys suggest areas of strength and weakness in survey methodology as recently applied to medical practice.

Methods

We reviewed published articles and publicly available documentation on the USC studies (*11,12*), NAMCS (*13,14*), and MEDECON (*15*), and computed a series of summary statistics from the USC and NAMCS data for the purpose of comparing results.

The basic features of the USC, NAMCS, and MEDECON studies are outlined in the box. In all three, physicians were selected from the AMA Masterfile (NAMCS used the American Osteopathic Association Masterfile as well). The USC studies took place between 1976 and 1978; NAMCS and MEDECON are continuing studies. USC requested the greatest volume of information from individual respondents, asking them to complete extensive questionnaires about the characteristics of their practices, keep logs of telephone contacts with patients, and record time spent for all purposes during assigned weeks. NAMCS included a shorter practice questionnaire and less detailed log diary for face-to-face patient encounters than USC. MEDECON requested completion of a questionnaire on practice characteristics.

We drew upon a study by Perrin and coworkers (*16*) to evaluate the reliability of key items in the USC surveys. In this study, a sample was selected of approximately 600 physicians who had earlier completed the USC data collection instrument. Contacting physicians 13 weeks after the most recent approach by USC, Perrin and coworkers obtained reliability data on 219 physicians through mailed instruments and on 374 through telephone interviews. Reliability of selected items was evaluated on the basis of compatibility between responses obtained initially by USC and during followup on the basis of two statistics: (*a*) kappa, a statistic reflecting the comparability of nominal

Table 1. Reliability of selected University of Southern California (USC) survey items

Items	Reliability indices		
	Kappa	Proportion of agreement	Spearman coefficients
Physician profile (from questionnaire)			
Primary specialty	.73	...	...
Primary practice arrangement	.65	...	...
Locations where patients are seen:			
Office	.35	.96	...
Outpatient department	.24	.76	...
Clinic	.23	.77	...
Hospital	.35	.88	...
Emergency room	.19	.65	...
Nursing home	.75	.87	...
Industry or school	.47	.90	...
Home	−.01	.84	...
Total number of office staff[1]	...	...	.58–84
Outpatients seen per week	...	...	.80–.91
Hours worked per week	...	...	.60–.78
Patient encounters (from log diary)			
Seen patient before?	.60	.93	...
Regular patient	.52	.89	...
Patient source	.37	.95	...
Diagnosis:			
Hypertension	.61	.96	...
Pharyngitis	.55	.96	...
Upper respiratory infection	.39	.94	...
Diagnostic procedures:			
Routine laboratory test	.56	.86	...
Blood chemistry	.26	.95	...
Culture	.58	.95	...
Chest X-ray	.33	.96	...
Other radiology	.51	.97	...
Therapeutic procedures:			
Immunizations	.64	.95	...
Injections—other	.28	.83	...
Patient education	.12	.91	...
Listening, reassurance	.27	.91	...
Systemic drugs	.32	.67	...
Topical drugs	.33	.95	...
Exercise, diet	.08	.91	...
Counseling: treatment program	.01	.90	...
Referral, consult	.33	.91	...

[1] Includes RNs, laboratory and X-ray technicians.

variables, controlling for marginal distribution; and (*b*) Spearman correlation, a statistic reflecting the comparability of continuous, ordinal variables.

We compared results from USC and NAMCS by focusing on cases and data elements falling within the scope of both surveys. This decision required dropping a large number of cases from the comparison that, although they were included in the USC sampling frame, were excluded from that of NAMCS. Comparison was restricted to those USC-surveyed physicians who provided private (non-Federal), office-based care within the 48 contiguous States.

Findings

Survey content and physician response rate. NAMCS, USC, and MEDECON obtained very different rates of response. Approximately 78 percent of the physicians sampled responded to the 1977 NAMCS survey, 57 percent to the USC surveys, and 40 percent to the MEDECON survey of 1978. The response rate of 57 percent represents an overall average of 24 separate surveys of medical and surgical specialties conducted by USC. Response rates to these surveys varied from a high of 82 percent for endocrinology to a low of 34 percent for general surgery:

Survey	Response rate (percent)
NAMCS (1977)	78
USC (total)	57
Response rates by specialty (examples):	
Endocrinology	82
Rheumatology	78
Nephrology	69
General internal medicine	53
Pediatrics	50
Family practice	44
General practice	36
General surgery	34
MEDECON	40

MEDECON, the survey with the least extensive set of questions and the most compact research instrument, obtained the lowest overall response rate.

Reliability of survey items. Table 1 summarizes results of Perrin and coworkers' study of the reliability of USC data elements. The level of kappa generally accepted as indicating reliability is around 0.60 (*17*). Since Spearman's rho is comparable to kappa for continuous variables, the same magnitude is regarded as a criterion of reliability. Reliability of three key, continuous variables appears acceptable. Spearman coefficients presented for the number of office staff, outpatients seen per week, and hours worked per week range from 0.58 to 0.91. The reliability of a majority of the items, whose kappas and rhos are well below 0.60, is questionable.

Table 2. Relative frequency of primary diagnoses in nonhospital settings encountered by general practitioners according to USC and NAMCS data[1]

USC study		NAMCS study	
Condition	Percentage of encounters	Condition	Percentage of encounters
Essential benign hypertension (401)	5.5	Essential benign hypertension (401)	5.9
Acute upper respiratory infection (465)	3.0	Acute upper respiratory infection (465)	4.4
Diabetes mellitus (250)	2.6	Ischemic heart disease (412)	2.6
Neurosis (300)	2.5	Diabetes mellitus (250)	2.5
Eczema and dermatitis (692)	2.3	Obesity (277)	2.4
Acute pharyngitis (462)	2.2	Eczema and dermatitis (692)	2.1
Ischemic heart disease (412)	2.1	Acute pharyngitis (462)	2.0
Obesity (277)	1.8	Cystitis (595)	1.8
Hypertensive heart disease (402)	1.7	Influenza with pneumonia (471)	1.8
Acute tonsillitis (463)	1.4	Neurosis (300)	1.7
Cystitis (595)	1.4	Influenza (470)	1.7
Diarrheal disease (9)	1.3	Acute tonsillitis (463)	1.6
Osteoarthritis (713)	1.2	Arthritis, unspecified (715)	1.4
Strains, sprains (847)	1.1	Otitis media (381)	1.4
Otitis media (381)	1.1	Osteoarthritis (713)	1.2

[1] Omits nondisease ICDA categories.

NOTE: Number in parentheses is ICDA code associated with diagnosis.

Replicability of survey findings. We compared three sets of findings from the USC surveys and NAMCS: (*a*) presenting conditions encountered by respondents, (*b*) characteristics of patients seen with two frequently encountered presenting conditions, and (*c*) actions taken by physicians in response to these conditions.

Table 2 compares the frequencies of primary presenting conditions encountered by general practitioners surveyed by USC in 1977 and by NAMCS in 1977 and 1978. Total frequencies for some of these diseases, including diabetes mellitus, acute tonsillitis, and osteoarthritis, were identical, or nearly identical, in both surveys. The greatest difference between USC and NAMCS appeared in the reporting of acute upper respiratory infections, USC respondents indicating that this disease represented 3.0 percent of the conditions they encountered, NAMCS respondents, 4.4.

Tables 3 and 4 compare USC and NAMCS findings on cases of essential benign hypertension seen by general internists and acute tonsillitis and pharyngitis seen by pediatricians. Table 3 indicates that the two surveys produced similar findings on the patient's sex and whether the responding physician had seen the patient before. Greater differences between USC and NAMCS, however, appear in the reporting of second diagnoses. Table 4 presents the greatest differences which are in the diagnostic and therapeutic procedures for the two conditions. Among the internists' encounters for essential benign hypertension, the two surveys produced similar findings only for use of the electrocardiogram. Among encounters between pediatricians and patients with tonsillitis and pharyngitis, only ordering of X-rays was reported about as frequently in the USC survey and NAMCS.

Table 3. Percentage of physicians reporting selected encounter characteristics for two conditions: comparison of USC and NAMCS data

Condition	Male patient	Second diagnosis present	Seen patient before
Essential benign hypertension (internists only):			
USC	40.8	40.0	94.2
NAMCS	39.5	53.2	94.9
Tonsillitis and pharyngitis (pediatricians only):			
USC	50.3	22.5	92.5
NAMCS	51.3	34.2	92.3

Evidence of both similarities and differences in findings from the two surveys emerges in interspecialty comparisons as well. According to NAMCS, internists spent an average of 18.4 minutes per encounter with patients with essential benign hypertension, as opposed to 13.0 minutes for family practitioners. According to USC, internists spent 17.9 minutes, family practitioners, 11.0. In responses to NAMCS, internists reported providing "counseling" in 19 percent of the visits for essential benign hypertension, as opposed to 12 percent among family practitioners. According to

Table 4. Percentage of physicians delivering selected components of care for two conditions: comparison of USC and NAMCS data

Condition	X-ray	Laboratory	Systemic drugs	Injections	Electrocardiogram
Essential benign hypertension (internists only):					
USC	14.4	25.4	73.3	1.8	15.7
NAMCS	7.4	28.6	53.7	4.6	15.1
Tonsillitis and pharyngitis (pediatricians only):					
USC	.6	64.8	66.0	16.4	. . .
NAMCS	.5	35.6	74.9	20.1	. . .

USC surveys, internists provide this service 23 percent of the time, and family practitioners, 20 percent. Interspecialty differences in time per visit indicated by the two surveys are similar in magnitude. The magnitudes of interspecialty differences in frequency of providing counseling inferrable from USC and NAMCS are not nearly as similar. Still, both NAMCS and USC indicate that internists report counseling more often than family practitioners.

Discussion

Clearly, many physicians are willing to complete lengthy research instruments. Less burdensome instruments do not necessarily result in higher response rates. NAMCS, for example, required respondents to report more information than MEDECON, but drew a considerably higher response rate. The more successful studies in this respect emphasize personal contact. While MEDECON approached physicians via letters from the editors of Medical Economics, NAMCS sent representatives to visit potential respondents. USC sent letters from national and local leaders of the medical community and used networks of physicians in specific localities to strengthen support for the study and included telephone followups of nonrespondents.

Ultimately, the success of a survey of physicians may depend most upon the degree to which potential respondents feel its results will promote better health care, and its sponsors share their perspectives and goals. Observations of rates and patterns of responses to serveral surveys support this proposition. Goodman and Jensen (*18*), for example, reported that response rates to the AMA's Periodic Survey of Physicians declined markedly between 1966 and 1977. They suggested that membership in the AMA was a strong factor in the decision to respond to the PSP. He noted that AMA members respond most often to the PSP, and that the percentage of physicians belonging to the AMA fell substantially during the 1960s and 1970s, particularly among younger practitioners. In contrast to the experience of PSP, rates of response to NAMCS—whose sponsors may project a more neutral image than the AMA—rose between 1973 and 1977 (*19*).

While surveys may produce adequate response rates, data such as those presented in table 1 raise questions about the reliability of their findings. The statistics reported in table 1, however, must be understood in the context of the reliability study that produced them. It was found that different individuals—physicians and office personnel, for example—often completed the questionnaire for the initial survey and reliability study, and that respondents to the followup study often relied on medical records that possibly contain errors.

The kappa statistic, moreoever, represents a conservative approach to analysis of reliability data. Because its computation formula adjusts for chance agreement, it underrepresents reliability of variables with highly skewed distributions, that is, those to which responses are likely to occur in one or a few categories. This consideration affects the vast majority of the variables represented in table 1; most are dichotomous with highly skewed distributions. Many variables with low kappas have high proportions of agreement. The item requesting patient source (self-referral, referral by another physician, and so forth, for instance, has a kappa of .37; the proportion of agreement of responses to the initial and followup survey, though, is 0.95.

After factors militating against measurable agreement on the initial and followup studies have been considered, the reliability estimates in table 1 appear more favorable. Those which appear the least reliable, moreoever, tend to be those measured by items whose response options are easily confused or interpreted by respondents in diverse ways. An item such as "systemic drugs," for

example, may be readily confused with "injections."

We apply only impressionistic criteria in assessing whether the USC and NAMCS findings replicate each other. The multifaceted sampling design of these surveys makes rigorous testing of differences between the two surveys' findings extremely burdensome. The variability of standard errors in five strata used in each USC study, for example, precludes the use of standard *t*-tests. To perform *t*-tests on each variable of interest, investigators must observe its variability within a series of randomly generated subsamples and produce an aggregate statistic (*20*). Performance of this procedure for more than a few variables greatly exceeds the scope of this paper, and no doubt the resources of most other investigations.

Table 2 presents findings for the two studies that appear fundamentally similar. The differences evident in the table may well be explained by relatively minor features of each study. The fact that USC surveyed general practitioners in July and August, while NAMCS surveyed equal numbers of physicians during each week of the year, could have resulted in the higher incidence of upper respiratory infection reported by NAMCS. Different coding routines could have distorted the reported frequencies of specific diseases. The relative unreliability of upper respiratory infection as a diagnosis detected in Perrin and coworkers' study further suggests that coding problems may have given rise to the USC–NAMCS discrepancy for this diagnosis.

Table 3 and 4 are more problematical. Still, differences between USC and NAMCS findings in these tables may reflect differences in technical details of the studies rather than problems with the overall methodology. USC, for example, provided many more response options of content of care than NAMCS, a factor that may help explain the relatively large differences observed in table 4, where most of the major differences occur. Coding of presenting conditions may also have differed in the two investigations. Sampling error may well explain the other differences in tables 3 and 4.

Conclusions

The data presented previously suggest that the large-scale survey can provide useful information about medical pratice. The quality of data obtained through this method, though, depends on the manner with which it is applied. The ultimate value of data obtained in this way, moreover, may depend on appreciation of limitations which cannot be ruled out in the large-scale survey.

'Investigators contemplating surveys of medical practice should assess their resources in this area, including their capacity for making personal contact with each potential respondent before confronting him or her with a research instrument.'

Response rates approximating those of NAMCS and the more successful USC survey efforts are clearly desirable and necessary for many applications. The experience of NAMCS and USC indicates that researchers may obtain high response rates even from busy clinicians. But only the confidence of the target population and considerable effort in making contact with potential respondents can ensure an adequate response rate. Investigators contemplating surveys of medical practice should assess their resources in this area, including their capacity for making personal contact with each potential respondent before confronting him or her with a research instrument. In view of differing values and affiliations among physicians, researchers should identify likely response bias and accordingly limit the generality of statements based on individual surveys.

Reliable survey items are necessary to allow meaningful interpretation of responses whatever the rate at which they are obtained. Obtaining reliable data from respondents, in turn, requires adherence to two basic rules of survey research: (*a*) making questions clear and (*b*) asking only questions that respondents have the resources to answer (*21*). Several items in the surveys discussed previously appear deficient in these areas. The data presented in this article suggest that imprecision in survey items contributes to differences in both initial and followup responses from USC respondents and in responses to similar items in the NAMCS and USC surveys. Reliability data such as those presented for the USC studies are seldom available. But users of data from other studies should acknowledge the possibility of imprecision in survey items not conforming to the two rules stated previously.

Nevertheless, data from a given survey of medical practice may retain much of their value to

scholars and policymakers even if they are not clearly replicated by other surveys. Without constituting proof of replicability, the USC-NAMCS comparison supports the validity of interspecialty comparisons made on the basis of large-scale surveys. The data we have presented suggest that USC and NAMCS consistently indicate interspecialty differences in the same direction. Internists, for example, reported longer encounters than family practitioners with patients who have diagnosed essential benign hypertension on both NAMCS and the USC surveys—although average encounter times for each specialty differ in USC and NAMCS.

Estimates of population parameters (for example, percentage of patients with specific diseases seen by individual specialties, frequency with which specific components of care are used) are more problematical. Recognizing the possibility of unknown biases in any survey of medical practice, those who require parameter estimates for scholarly or policy-related purposes may limit themselves to drawing inferences supportable only on the basis of ranges of estimates obtained from several surveys. Lubitz recommends a similar strategy for users of various national hospital utilization surveys, which have produced significantly different findings (*22*). Alternatively, survey users may select the data files emerging from the study with the clearest biases and sets of response options most compatible with their specific needs, recognizing, of course, attendant limitations in their analyses.

References

1. Reinhardt, U.: Physician productivity and the demand for health manpower. Ballinger Publishing Co., Cambridge, MA, 1975.
2. Sloan, F. A.: A microanalysis of physicians' hours of work decisions. *In* The economics of health and medical care, edited by M. Perlman. John Wiley and Sons, New York, 1973.
3. Aiken, L. H., et al.: The contribution of specialists to the delivery of primary care. N Engl J Med 300: 1363-1370, June 14, 1979.
4. Noren, J., Frazier, T., Altman, I., and DeLozier, J.: Ambulatory medical care: a comparison of internists and family-general practitioners. N Engl J Med 302: 11-16, Jan. 3, 1980.
5. Fishbane, S., and Starfield, B.: Child health care in the United States. N Engl J Med 305: 552-556, Sept. 3, 1981.
6. Rosenblatt, R. A., Cherkin, D. C., Schneeweiss, R., and Hart, L. G.: The content of ambulatory medical care in the United States, N. Engl J Med 309: 892-897, Oct. 13, 1983.
7. McNutt, D. R.: GMENAC: its manpower forecasting framework. Am J Public Health 71: 1116-1124 (1981).
8. Peterson, O. L.: An analytical study of North Carolina general practice, 1953-1954. Association of American Medical Colleges, Evanston, IL, 1956.
9. Clute, K. F.: The general practitioner: a study of medical education and practice in Ontario and Nova Scotia. University of Toronto Press, Toronto, 1963.
10. Marsland, D. W., Wood, M., and Mayo, F.: A data bank for patient care, curriculum, and research in family practice: 526,196 patient problems. J Fam Practice 3: 25-28 (1976).
11. Girard, R. A., et al.: A national survey of internal medicine and its specialties. I. An overview of the practice of internal medicine. Ann Intern Med 90: 965-975 (1978).
12. Mendenhall, R. C., et al.: A national study of medical and surgical specialties: II. Description of the survey instrument. JAMA 240: 1160-1168, Sept. 8, 1978.
13. National Center for Health Statistics: The National Ambulatory Medical Care Survey: background and methodology. DHHS Publication No. (HRA) 76-1335. U.S. Government Printing Office, Washington, DC, 1974.
14. National Center for Health Statistics: Patterns of ambulatory care in general and family practice: the National Ambulatory Care Survey. DHHS Publication No. (PHS) 83-1734. U.S. Government Printing Office, Washington, DC, 1983.
15. Owens, A.: Could you justify higher fees now? Med Economics 54: 211-223 (1977).
16. Perrin, E. B., Harkins, E. B., and Marini, M. M.: Evaluation of the reliability and validity of data collection in the USC medical activities and manpower projects. Final report. Health and Population Study Center, Battelle Human Affairs Research Centers, Seattle, WA, 1978.
17. McKennell, A.: Attitude measurement: uses of the coefficient alpha with cluster or factor analysis. Sociology 4: 227-245 (1970).
18. Goodman, L. J., and Jensen, E.: Economic surveys of medical practice: AMA's periodic surveys of physicians, 1966-1978. *In* National Center for Health Services Research: Health survey research methods: third biennial conference. DHHS Publication No. (PHS) 81-3268. U.S. Government Printing Office, Washington, DC, 1981, pp. 56-65.
19. Loft, J. D.: Methodology of the National Ambulatory Medical Care Survey: evaluation and extensions. *In* National Center for Health Services Research: Health survey research methods: third biennial conference. DHHS Publication No. (PHS) 81-3268. U.S. Government Printing Office, Washington, DC, 1981, pp. 68-78.
20. Mosteller, F., and Tukey, J. F.: Data analysis and regression. Addison-Wesley Publishing Company, Inc., Reading, MA, 1977.
21. Babbie, E. A.: Survey research methods. Wadsworth Publishing Co., Belmont, MA, 1973.
22. Lubitz, J.: Different data systems, different conclusions? Comparing hospital use data for the aged from four data systems. Health Care Financing Rev 2: 41-60 (1981).