

Care After the Onset of Serious Illness: A Novel Claims-Based Dataset Exploiting Substantial Cross-Set Linkages to Study End-of-Life Care

NICHOLAS A. CHRISTAKIS, M.D., Ph.D., M.P.H.,¹
THEODORE J. IWASHYNA, M.D., Ph.D.,² and JAMES X. ZHANG, Ph.D.³

ABSTRACT

To date, there has not been a study using a large, nationally representative group of patients with serious illness who are at risk for hospice use and who are followed forward in time to understand the determinants of hospice use. In this paper, we outline the development of a large new cohort of 1,221,153 Medicare beneficiaries newly diagnosed with 1 of 13 serious conditions in 1993, a cohort that can be used to study end-of-life care in the United States. In describing our methods, we illustrate the possible utility of Medicare claims for end-of-life research. The members of our cohort are followed forward for hospice and other health care use through December 1997, and for mortality through June 1999. Medicare claims data on their inpatient and outpatient hospital use is also collected. Based on the ZIP Codes and counties in which cohort members lived, we were also able to characterize the health care markets of cohort members, as well as obtain other socioeconomic information about them. Information about cohort member's health care providers is also available. Detailed health information about cohort members' spouses was also collected. We conclude by highlighting the types of analyses that can be conducted in this data set.

INTRODUCTION AND OVERVIEW

EVERY YEAR, 2.3 million Americans die, many after a chronic disease of several years' duration.¹ More than 80% of Americans will die in an institution,² and almost as many will die in pain.³ The terminal months of life can be devastatingly expensive, both for families⁴ and for society.⁵⁻⁷ The foregoing observations have led to substantial discussion of ways in which end-of-life care might be reformed; this discussion has occurred

in professional,⁸ governmental,⁹ and lay forums.¹⁰ However, such discussions have, we believe, been limited by the absence of adequate data on the patterns of health care utilization among the seriously ill—that patient population for which prospective interventions may be effective. Certainly there have been some pioneering studies, ranging from the SUPPORT study³ to claims-based studies of total expenditures^{5,6} and of those who use hospice.¹¹⁻¹³ Similarly there have been nationally representative studies of the

¹Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts.

²Department of Medicine, Hospital of the University at Pennsylvania, Philadelphia, Pennsylvania, and Harris Graduate School of Public Policy Studies, University of Chicago, Chicago, Illinois.

³Department of Community and Family Medicine, School of Public Health, The Chinese University of Hong Kong, Hong Kong.

elderly, such as the Longitudinal Study on Aging,¹⁴ the Health and Retirement Survey,¹⁵ the Panel Study on Income Dynamics,¹⁶ and the Medicare Current Beneficiary Survey.¹⁷ There have also been long-term studies following the disablement process, such as the National Long-Term Care Survey. However, to the best of our knowledge, no one has examined a nationally representative cohort of elderly individuals from the point of their diagnosis with serious illness until death. The goal of such a study would be to allow denominator-based comparisons of the roles of individual clinical and demographic factors, as well market structure, on the timing and nature of end-of-life care.

Here we describe the development of the Care after the Onset of Serious Illness (COSI) dataset. This dataset seeks to explore the potential for multiple levels of analysis made possible by the existence of many different electronic datasets. These data include Medicare claims data regarding patients' inpatient and outpatient health care use and regarding their medical providers, as well as other data sources such as the Census, American Hospital Association data, and the like.

An overview of the entire data extraction and linkage process is presented in Figure 1. This paper also serves as a record of the rationale, both empiric and theoretic, for a number of the design decisions that must be made in the development of such a data set. Finally, we provide an illustration of the types of analyses that are possible. In a companion paper, we provide some disease-specific rates regarding hospice and other health care use.¹⁸

EMPANELMENT

General data source

The core data used to develop COSI are the 1993 inpatient hospitalization records from the Health Care Financing Administration's (HCFA) Medicare program. These records, embodied in the "MedPAR" file, represent a complete enumeration of the final adjudication of all claims for hospitalizations filed by all Medicare beneficiaries for any hospitalizations or parts thereof occurring at any time during 1993. Medicare claims

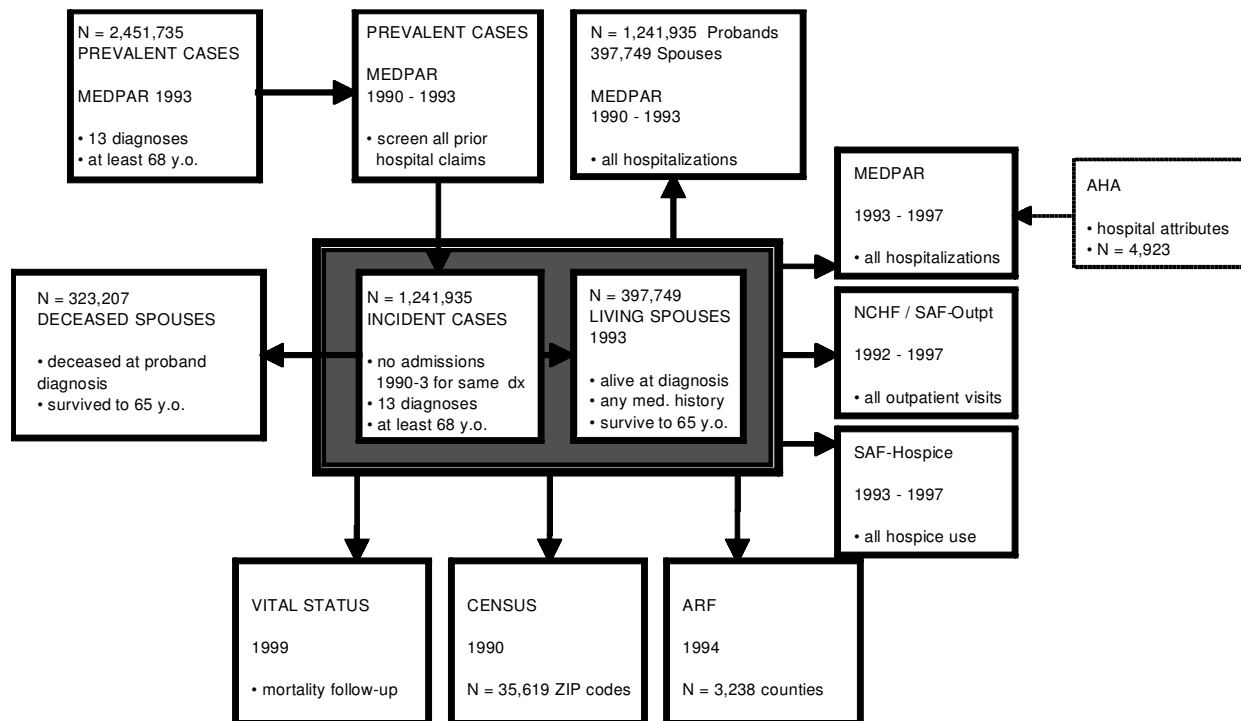


FIG. 1. Overview of Data Linkage. MedPAR, Medicare Provider Analysis and Review File (Inpatient Claims); Dx, diagnosis; y.o., years old; AHA, American Hospital Association; NCHF, National Claims History File (Part B Outpatient Claims to Physicians and Suppliers); SAF-Outpt, Standard Analytic File for Outpatient Claims; SAF-Hospice, Standard Analytic File for Hospice; ARF, Area Resource File.

TABLE 1. CONCEPTUAL CRITERIA FOR SELECTING DISEASES FOR INCLUSION IN THE COSI COHORT

Severity: the disease had a reasonably high probability of substantial mortality or morbidity
Acuteness of onset: the natural history of the disease is marked by a point of threshold increase in its manifestations
Likelihood of hospitalization: the threshold increase in the manifestation of the disease is <i>very</i> likely to result in an acute hospitalization regardless of other characteristics of the patient
Reliability of detection: the disease needed to be detectable in the claims with both high sensitivity and specificity
Epidemiologic significance: the disease must account for a reasonable burden of disease in the population
Theoretical interest: the diseases should be sufficiently different in their natural histories so as to allow generalizations in other domains.

COSI, Care After Onset of Serious Illness.

and enrollment data capture 96% of the American 65-plus population.¹⁹

Selection of initial conditions and their operationalization

The COSI project focuses on the longitudinal course of patients who have the new onset of a serious disease in 1993. The year was chosen arbitrarily, far enough back that substantial follow-up would be available, sufficiently recent so as to take advantage of Medicare's significant improvements in the quality of its electronic records at the end of the 1980s. Our objective was to construct an inception cohort of patients newly diagnosed with one of several serious illnesses based on examination of hospital records. We established a number of conceptual criteria for diseases to be included in the cohort.²¹ (Table 1). Thirteen diseases were selected that met these criteria^a (Table 2).

To define cases, we relied on International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes in the hospital claims; this required selecting ICD-9-CM codes that represent these diseases. In many cases, multiple definitions were identified in the published

literature (Table 2). We then chose the definitions with the best published empirical performance, when available. In the absence of comparative performance data, we chose definitions that most coincided with the best accepted research in the subfield, for example, the Surveillance, Epidemiology, and End Results (SEER) definition for cancers, or the Cooperative Cardiac Project definition for acute myocardial infarction. Table 2 provides the actual definitions, as well as references and comments.

After defining the ICD-9-CM codes to be used to identify a diagnosis, the appropriate exclusions needed to be implemented in order to capture incident cases. Prior detailed empirical work provided guidance here. Research examining the Medicare/SEER linked data demonstrated that for lung, colon, and esophageal cancers, 3 years of look-back in the Medicare claims was adequate to eliminate prevalent cases.²⁰ That is, if an individual had not been hospitalized in the prior 3 years before the putative index hospitalization for onset of his/her serious disease, it was very likely that they had never previously been hospitalized for the disease. As such, hospitalizations for a disease with no similar hospitalizations in the past 3 years served as our operational definition of an incident index hospitalization.

Hospital claims records can have up to 11 diagnoses, 1 in the "primary position" for the disease most responsible for the hospitalization, and 10 in "secondary positions" for diseases that contributed to the stay. Once the look-back had been defined, it was necessary to decide whether (1) to require that onset of a disease be defined as only those hospitalizations for which the disease of interest was noted to be the primary cause of hospitalization, or (2) to also accept as index hospitalizations those hospitalizations where the disease was noted to be contributory to the pa-

^aBreast cancer, prostate cancer, chronic obstructive pulmonary disease, dementia, and still other diseases and conditions were considered as other COSI-empanelling diagnoses. Breast and prostate cancer were excluded based on the SEER/Medicare evidence that an accurate index hospitalization could not be identified in the sense that (1) not all patients with the disease are hospitalized for this condition, and (2) when they are, the hospitalization is not usually very near the diagnosis.²⁰ Clinical experience suggested that the natural histories of C.O.P.D. and dementia were stories of gradual worsening with occasional hospitalizations only in some situations; moreover, claims data have been shown to be problematic in dementia.

TABLE 2. ICD-9-CM OPERATIONALIZATIONS OF COSI DIAGNOSES

	ICD-9 codes	Source	Others considered
Noncancer			
MI	410.0–410.9 (exclude 410.x2)	(22, 109)	(110–112)
CHF	398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 428.0–428.9	(113)	(45, 110–112, 114)
Hip fracture	820–820.9	(112)	(115)
Stroke	434, 436	(116, 117)	(110–113, 118, 119) ^b
Cancer			
Colorectal	153–154.8	(112)	(20) used colon only
Lung	162.2–162.9	(20, 120)	
Urinary ^a	188–189	(121)	
CNS	191, 192 194.3, 194.4	our own	(122–124) used intracranial tumors only
Head/neck	140–149, 161	our own	(121, 125, 126) were combined for this project.
Leukemia	204–208.9	(121, 123, 127)	
Lymphoma	200–203; 238.6	(121)	
Liver/biliary	155–156	(121)	
Pancreatic	157	(121)	

^aUrinary tract cancers do not include prostate cancers.

^bA comparison of multiple definitions with chart review is available.¹¹⁶

MI, myocardial infarction; CHF, congestive heart failure; CNS, central nervous system; ICD-9-CM, International Statistical Classification of Diseases, 9th revision, Clinical Modification; COSI, Care After the Onset of Serious Illness.

tient's hospitalization. As Table 3 demonstrates, these differences could lead to substantial differences in the apparent incidence of the disease. Here, we relied on three sources: precedents, alternative epidemiologic data, and clinical experience. Table 4 demonstrates the comparison between our final definitions and existing epidemiologic data based on sources other than the Medicare claims, where such data exist.

For cancer patients, we accepted as the index hospitalization any hospitalization during 1993 that indicated a cancer diagnosis as defined in any position, as long as the patient had never previously had a hospitalization where this cancer diagnosis had been noted. Naturally, the exclusion criteria were disease-specific; a diagnosis of acute myocardial infarction in 1992 did not prevent a patient newly hospitalized in 1993 with a lung cancer diagnosis from entering our cohort. A similar line of reasoning went into the choice of methods for congestive heart failure (CHF).^b

^bIn the sole case of CHF, because of the commonness of the disease, we took a simple random 1 in 3 sample of all detected CHF patients to enter into our study cohort.

Unlike most cancers or CHF, it is quite possible for a patient to have more than one incident stroke, myocardial infarction (MI), or hip fracture. Therefore, the use of a look-back to exclude prevalent cases is less satisfying—it may lead to an inappropriately healthy selection bias by excluding those with multiple cardiac or intracranial events or fractures. However, for stroke and MI, our clinical experience and the past practice of other researchers both agreed that for an individual having an incident event, that event would be their primary diagnosis for that index hospitalization. This choice is reinforced by the distinction in the ICD-9-CM, explored by the Cooperative Cardiovascular Project, between initial visits for an acute MI and follow-up care visits.²² In the case of hip fractures, it was not necessary to restrict the definition to only those cases where the diagnosis was in the primary position.

Cohort construction to this point, as represented in Table 3, has allowed individuals having index hospitalizations in more than one disease within 1993 that met our enrollment criteria to be enrolled multiple times. This was done to allow complete enumerations during disease-specific analyses to be

TABLE 3. ALTERNATIVE DEFINITIONS OF INCIDENCE: UNIQUE INDEX HOSPITALIZATIONS

	1° Position		Any position	
	Prevalent	Incident	Prevalent	Incident
Noncancer				
MI	256,183	234,098	323,736	296,144
CHF	494,845	299,161	1,294,707	833,027
				(sample = 277,676)
Hip fracture	218,729	207,927	228,677	216,431
Stroke	268,222	241,479	334,016	300,093
Cancer				
Colorectal	78,189	72,165	98,877	84,093
Lung	58,077	51,072	110,243	87,619
Urinary tract	39,553	31,142	54,964	40,897
"Bad" cancers:				
Leukemia	9,505	7,168	34,940	22,017
Lymphoma	22,182	16,671	53,042	34,327
Pancreatic	12,834	11,661	19,233	16,225
Liver/biliary	7,383	6,695	11,290	9,655
CNS	5,636	5,103	7,230	6,276
Head/neck	9,051	7,848	14,127	11,428

Bold numbers indicate the choice made for this project. This is based on acute inpatient hospitalizations with age at least 68 for any hospitalization (that is, without any geographic restrictions, claims completeness, or date validity checks).

MI, myocardial infarction; CHF, congestive heart failure; CNS, central nervous system.

maintained. During full-cohort analyses, patients were entered into COSI under temporally first diagnosis. That is, if a patient had a hospitalization in February 1993 that met criteria for an index hospitalization for MI, and then had a second hospitalization in May 1993 that met criteria for pancreatic cancer, that individual would appear in both the MI and pancreatic cancer subgroups, but would appear only once in the unified COSI analytic cohort, under an MI diagnosis. This occurred in 7.9% of unique individuals (Table 5).^c

Other empanelment issues: exclusions based only on inpatient acute care hospital claims

While the bulk of MedPAR claims are for acute hospitalizations (11,307,844 of 12,709,289 or 96%), a number of other types of "hospitals" are in-

cluded in the claims. These include primarily skilled nursing facilities, but also some so-called "long-term acute care facilities." We use only claims from acute hospitalization for empanelment index hospitalizations. This allowed greater consistency with previous work, and agreement with our model of what an index hospitalization "should" be. Similarly, claims for outpatient care (either Part A or Part B) were not used to detect the onset of disease, nor were they used to exclude patients as having prevalent rather than incident disease. In our judgment, we were interested in studying processes of care after the threshold change in disease intensity that our clinical experience suggests marks the first hospitalization for the diseases in COSI. To maintain consistency, a patient with a previous hospitalization at a non-acute care hospital or a previous outpatient visit for a COSI-defining diagnosis could still be included in the cohort; that is, only acute care hospitalizations were used to exclude prevalent cases.

Other empanelment issues: minimal data completeness restrictions

In order to be empanelled into our cohort, we required the following minimal data integrity

^cIn cases where the diagnoses from a single day that qualified a patient simultaneously for more than one diagnosis, the patient was empanelled separately into each of the disease-specific cohorts. They were empanelled into the combined cohort under the cancer diagnosis preferentially if there was a cancer and a noncancer that both qualified simultaneously. A total of 54,037 (4.2%) patients qualified for 2 subcohorts, 495 (0.04%) qualified for 3 on the same day, and 2 patients qualified for 4 cohorts on the same day.

TABLE 4. INCIDENCE OF CANCER DIAGNOSES IN THE ELDERLY IN COSI AND SEER DATABASES, IN THOUSANDS OF EVENTS PER YEAR

	SEER	1993 MedPAR ^a	COSI ^b
Colorectal	91	95	84
Lung	125	104	86
Leukemia	16	24.5	22
Lymphoma	38	39	34
Pancreatic	20	18	16
Liver/biliary	12.5	11.6	10
CNS	6.3	8.8	7.4
Head/neck	14.8	14.3	11.6
Urinary tract	59	47	41

^a1993 MedPAR, any age, no look-back to exclude incident cases.

^bCOSI data use for individuals older than 68 whereas SEER data do not have this restriction.

Source: SEER data is from Ries LA, Gloeckler CL, Kosary BF, Hankey BA, Miller AH, Edwards BK: 1997. *SEER Cancer Statistics Review, 1973–1994*. Bethesda, MD: National Cancer Institute. NIH Pub. No. 97-2789. Other columns are from authors' own tabulations. CNS, central nervous system.

checks from the claims: (1) valid birth date, in order to impose the age restrictions; (2) a valid admission date; (3) some valid ICD-9-CM codes that met our enrollment criteria. The presence of other data errors on a claim (e.g., some invalid ICD-9-CM codes in other diagnostic fields) did not exclude a claim from empanelment. Therefore, there were some remaining data errors in the claims that required exclusion of claims because of incoherent dates (e.g., death dates reported later than the day on which we took delivery of final mortality follow-up data or death dates before the admission date), missing gender, or missing race. The final cohort size with adequate minimum data was 1,231,894 unique individuals; this represents 99.19% of the 1,241,935 unique individuals initially screened for possible inclusion in cohort (i.e., 0.8% were excluded because of miscellaneous data impurities).

TABLE 5. NUMBER OF INDEX HOSPITALIZATIONS IN 1993 BY UNIQUE INDIVIDUALS

Count	Frequency	Percent of individuals
1	1,144,365	92.1%
2	93,892	7.6%
3	3,568	0.3%
4	109	0.0%
5	0	0.0%
6	1	0.0%

This is based on acute inpatient hospitalizations with age at least 68 for any hospitalization (that is, without any geographic restrictions, claims completeness, or date validity checks).

*Other empanelment issues:
geographic restrictions*

At this point we have not imposed any geographic restrictions. Thus, the cohort includes individuals in Puerto Rico, Guam, and the U.S. Virgin Islands, and other miscellaneous territories. As supplementary material such as Census data is not in general available for these areas, we excluded such cases from many analyses; 1,221,153 probands (99.1%) lived within the 50 United States and the District of Columbia.

FINDING SPOUSES

We have previously published a method to detect the marital status of many Medicare beneficiaries based on information latent in their claims. This method allows us to uniquely link individuals to their spouse. The details of this method have been described^{23,24} and discussed elsewhere.^{25,26} Briefly, it is has long been known that some married and widowed individuals file Medicare claims under a Health Insurance Claim number (HIC) that consists of their spouse's Social Security number and a code indicating that the filing individual is a “dependent beneficiary” rather than a primary beneficiary. Moreover, individuals can change the HICs they use, particularly when their spouse dies. This is one of the reasons for the use of “cross-reference files” when linking an individual across multiple years or types of HCFA data. What we noted previously

was that, while this causes hassles for constructing individual-level longitudinal data sets,²⁷ this also permits the development—using only information present in the claims—of longitudinal couple-level data sets.

In order to apply these methods, we could not restrict ourselves exclusively to the MedPAR files and other utilization-based claims. After all, a spouse might exist but not use any health care during our year of interest. However, HCFA also maintains an enrollment database. The 1993 Denominator file contains basic identifying information on the entire Medicare population during 1993, that is, it contains information on all individuals who were enrolled in Medicare at any point in 1993, regardless of whether or not they actually filed a claim. The enrolled population has been previously shown to closely approximate the population of all Americans age 65 and older.^{19,28}

We received from HCFA the 1993 Denominator file of 38,212,735 records, with mortality follow-up for the entire 1993 Medicare population through July 6, 1999, and a cross-reference file as of January 6, 1999. These mortality and cross-reference files were the most recent available at the time of this particular data request. We used this data to develop a list of all detectable husband-wife pairs as of 1993, where both were enrolled in Medicare at some point during that year. After doing so, we would “look-up” the spouse of our COSI-cohort members, matching them if possible. This allowed us to determine who was married (so far as we could detect) on January 1, 1993. We then took the unmatched population (i.e., the probands for whom we did not find a living spouse), and probed the Medicare data to see if they ever had had a spouse, which would allow us to not merely designate them a widow or widower, but to know for precisely how long they had been in such a status.^{29,30}

The 1993 Denominator file contains 38,209,888 unique individuals, 32,180,588 of whom were at least age 65 as of January 1, 1993. The results of the matching process are schematized in Figure 2. The final distribution of our cohort by marital status as of January 1, 1993 is shown in Table 6. Because the dates of the transitions from married to widowed are known precisely, these dates are recorded in the data set. Therefore, for cohort members with multiple index diagnoses, we can determine their precise marital status on the day of admission for the disease under consideration at that time in disease-specific analyses. More-

over, we can explore in detail the time course of effects of marital status transitions.

We cannot, using these methods, detect new marriages among the elderly, which is a very rare phenomenon. (In 1990, among the previously widowed, the annual rate was 1.7 marriages per 1000 for elderly women, and 14.0 per 1000 for elderly men; thus, it is unlikely that more than a few percent of our widowed sample may have remarried.^d Likewise, we cannot detect cohabiting couples, a similarly rare phenomenon at this point in time in this population.^{33,34}

DEFINING THE DEATH DATE

Death dates were obtained from the highly accurate Vital Status file of the Health Care Administration as of July 6, 1999. This file is updated regularly from the Social Security Administration. This file has been shown to be highly accurate, although there are known defects in the detection of death of certain old widows.²⁸

EXPANDING THE RACIAL AND ETHNIC CODING BASED ON NAME ALGORITHMS FOR HISPANICITY AND ASIAN ORIGIN

Medicare data have certain well-known limitations with respect to their racial classification system, and the race codes provided in the claims can only be reliably used for white/nonwhite comparisons.^{35,36} However, it is possible to apply

^dThese data may be contaminated by couples who are divorced. Some of the members of these former couples may qualify for dependent spousal benefits, although the restrictions are quite strenuous. While fewer than 5.7% of the elderly are divorced,³¹ some of these former couples may contribute to the overestimation of our efficacy at detecting spouses. There are two approaches to dealing with this that can be implemented. First of all, we can require that married couples have the same mailing address ZIP code at the time of the proband's admission, which occurs in 87.2% of cases. Second, so-called unequated BICs are present in files such as the Physician/Supplier and Outpatient claims for probands or spouses present in these claims. Using unequated BICs, we can note that the spouse is receiving benefits as a divorcee, and change their marital status appropriately; less than 2% of couples are detected as divorced using the 1992 and 1993 data from these files. Although the decision is application-specific, we typically require that couples be coresiding in order to be used as married in our analyses.

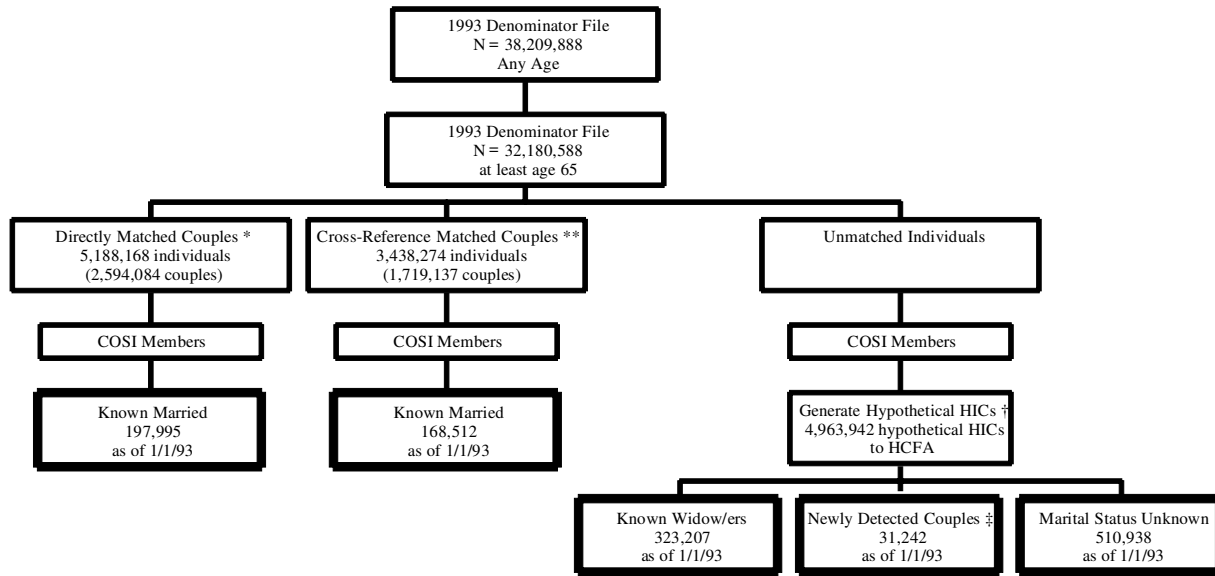


FIG. 2. Detection of spouses. *As was known, a single primary beneficiary may have more than one dependent beneficiary; this occurred 36,646 times. **Altogether, there were a total of 36,915,227 Health Insurance Claim numbers (HICs) in the cross-reference file used by the 32 million individuals ages 65-plus in the Denominator File. Of those, 10,110,008 have died by our follow-up on July 6, 1999. †In order to detect already deceased spouses, we took all the HICs ever used by any Care After the Onset of Serious Illness (COSI) cohort member not matched to a spouse alive in 1993 by the process using the Denominator file. We then changed the BICs to their reciprocal—where we had a BIC indicating a dependent spouse or widow, we created a HIC indicating a primary beneficiary, and vice versa. We obtained a Vital Status data—birth dates and death dates, if any—for all cases where the hypothetical HIC indicated an actual individual. ‡Newly detected married couples were those in which a spouse who had not died as of January 1993, was detected by the hypothetical HICs. These largely represent couples in which the non-COSI spouse qualified for Medicare after 1993, and so was not in the 1993 Denominator File, but before our hypothetical HIC file was processed in February 2000.

well-validated algorithms for identifying Hispanic and Asian-American ethnicities based on individual's names, substantially improving the adequacy of the racial/ethnic classification system we can use here.^{37,38} A total of 28,719 probands had their race codes reassigned by such a process. As expected given the geographical racial distribution,³⁹ 85.5% of these reclassified Hispanic and Asian-Americans lived in the states of Arizona, California, Florida, Hawaii, Illinois, Massachusetts, New Jersey, New Mexico, New York, and Texas.

DEVELOPING COMORBIDITY MEASURES

In order to make valid mortality comparisons between groups, differences in health at baseline must be taken into account. One fruitful way to operationalize "health" for such purposes is the notion of comorbidity burden. A comorbidity is a chronic disease of substantial mortality, mor-

bidity, or management burden. The number of such comorbid conditions affecting a patient is often computed to form a comorbidity index as a simple scalar measure.

Among the most popular comorbidity indices in claims data research are those based on the work of Charlson and collaborators,⁴⁰ particularly as implemented in the ICD-9-CM codes for computerized use.⁴¹⁻⁴³ While several alternative risk adjustment approaches have also been published,⁴⁴⁻⁵³ the Charlson method is extremely popular and has been used extensively.^{11,23,43,54,55} Direct comparisons between these alternative scales are relatively rare, and the choice of the Charlson index is somewhat arbitrary.⁵⁶⁻⁵⁹ On the whole, these indices have been developed for the prediction of mortality after hospitalization.

We have previously shown that statistically and empirically significant improvements in the prediction of mortality were obtained by incorporating alternative sources of data—particularly 2 years of inpatient lookback combined with 1

TABLE 6. MARITAL STATUS OF COSI COHORT MEMBERS AS OF JANUARY 1, 1993

	Total	%	Who died first?		
			Proband	Spouse	Neither
Men					
Method I-Directly Matched					
Proband is A	124,217	23.9%	24,334	73,376	26,507
Proband is B	3,304	0.6%	589	1,966	749
Method II-Cross-Reference Matched					
Proband is A	119,299	22.9%	22,161	89,094	8,044
Proband is B	3,427	0.7%	942	2,064	421
Follow-Up					
Proband is A	28,021	5.4%	1,500	18,595	7,926
Proband is B	263	0.1%	23	141	99
Widowed	54,242	10.4%			
Unmatched	187,049	36.0%			
Total	519,822				
Women					
Method I-Directly Matched					
Proband is A	1,907	0.3%	641	839	427
Proband is B	68,567	9.6%	25,195	29,410	13,962
Method II-Cross-Reference Matched					
Proband is A	5,387	0.8%	2,248	2,936	203
Proband is B	40,399	5.7%	16,942	17,959	5,498
Follow-Up					
Proband is A	654	0.1%	83	411	160
Proband is B	2,304	0.3%	833	983	488
Widowed	268,965	37.8%			
Unmatched	323,889	45.5%			
Total	712,072				

COSI, Care After the Onset of Serious Illness.

"Methods I," "Method II," and "Follow-Up" refer to ways at finding spouses. "A" and "B" refer to the BICs which identify whether an individual is a primary or secondary beneficiary. "Unmatched" means we were not able to find a spouse (alive or dead) for the proband.

year of outpatient and auxiliary claims look-back—but only if indices derived from distinct sources of data are entered into the regression distinctly.⁶⁰ Furthermore, we found that these improvements in explanatory power largely held regardless of whether one also controlled for Charlson scores based on self-reported health history and/or based on the secondary diagnoses from the claim for the index hospitalization. Therefore, for COSI, we computed separate Charlson scores for each data source for 1-year intervals prior to each index admission date; this means that for individuals empanelled with multiple diseases, they have multiple, diagnosis-specific Charlson scores. We have also computed the comorbidity adjustment approach proposed by Elixhauser et al.^{45 e}

^eWe are unaware of information comparing the performance of the Charlson score derived from multiple data sources and the Elixhauser score.

EXTERNAL DATA LINKAGES: INSTITUTIONAL PROVIDERS

Hospitals

There were a total of 5103 hospitals in the MedPAR data; not all of these hospitals were included in COSI. In 1993, 5084 had at least 10 Medicare discharges and could be identified in HCFA's Provider of Service File. From this information, we linked to the 1993 American Hospital Association (AHA) Annual Survey data.⁶¹ The AHA data is a survey of all hospitals; it is typically considered the best self-reported source of information on hospital features. Using hospital names, local address, and telephone number, from HCFA, we were able to link to a total of 4923 (96.8%) short-term acute care hospitals in the AHA annual survey database. This accounts for 91.4% (4923/5387) of all adult short-term acute care hospitals identified by the AHA. We also

identified all hospitals that were ranked “Best Hospitals” by the *U.S. News & World Report* annual ranking in 1993 in the following six specialties particularly germane to the care of COSI patients: cardiology, gastroenterology, neurology, oncology, orthopedics, and geriatrics.^{62,63} These hospitals may have improved outcomes relative to all other hospitals for patients with MI⁶⁴; the rankings are certainly highly visible.

Physicians

It would be possible to link our data to HCFA's MPIER file or the AMA Masterfile at the physician-level. For example, approximately 300,000 different physicians billed for outpatient services to the COSI patients.

EXTERNAL DATA LINKAGES: INDIVIDUAL PATIENTS

Overview

A major limitation of claims-based data explorations is the paucity of individual-level information about non-health-related attributes or outcomes. In the current project, we have attempted to overcome this in two ways. First, we have tried to maximally exploit the information available from HCFA, using the indicators of Medicaid receipt,⁶⁵⁻⁶⁷ the marriage detection algorithm, expanded ethnicity detection algorithms, and detailed comorbidity measures. Second, we have taken advantage of the many high-quality local area datasets available from the U.S. government; in particular, we linked to the 1990 Decennial Census and to the Area Resource File. While these auxiliary data sets do not provide individual-level detail, they provide important information about the communities in which our probands make their lives. For many studies, this local area information is quite useful.

U.S. Census

Data were linked to the 1990 U.S. Decennial Census; the Census provides the most detailed information about population characteristics available.⁶⁸ This was done at the ZIP-Code level. ZIP Codes are aggregations of 25,000 to 50,000 residents developed for administrative purposes. As such, they do not necessarily represent community boundaries in the way community areas or

census tracts attempt to. However, because of their ready availability and relatively low level of aggregation, they are often used in linkage studies.^{67,69-74} We were able to link 1,184,995 (97.1%) of the 1,221,153 probands who were in the 50 states and the District of Columbia to the 1990 Census. The linkage failures likely resulted from data errors in the claims and the Postal Service's periodic creation of new ZIP Codes in dense areas. From this, we were able to extract information about the communities in which the probands reside, such as the age distribution, race, median income, median education-level, and population density.

We were particularly interested in the use of Census data to provide additional information on the level of affluence of the communities in which our probands reside. This provides a continuous measure that is likely well-correlated with household-level total financial resources. The interpretive validity of this approach has been documented^{75,76}; however, there remain certain limitations as to the interpretation of any estimated effects from such proxy values.⁷⁷⁻⁷⁹ There is an extensive debate on the usefulness of such area-based measures in the literature.⁸⁰⁻⁸⁴ The major interpretive difficulty arises because geographical data may tend to undercontrol for variation in economic resources; for example, it will fail to take into account the fact that African Americans in general have lower levels of wealth at the same income levels as whites.⁸⁵ However, among the elderly, the use of area-measures may in fact better approximate the concept of mobilizable financial resources, such as home equity, than would a simple measure of income. For health decisions, particularly at the end-of-life, a more general measure of assets may be more appropriate for studying the influence of finances on choices. Furthermore, these aggregate measures are necessary to study many types of questions.⁸⁶⁻⁸⁸

Area Resource File: county-level definition of market variables

The Area Resource File (ARF) is a publicly available aggregation of data from a number of sources produced by the federal government. It is commonly used in health services research to provide information at the county level.^{69,89-97} We were able to link 1,203,919 (98.6%) of the 1,221,153 probands who resided in the 50 states and the District of Columbia. There was no par-

ticular relationship between whether patients could be linked to the Census via ZIP Codes or linked to the ARF via county information. From the ARF we could obtain measures of health care infrastructure, population demographics, and other variables.

Counties were particularly of interest to us as we wanted to define the health care markets relevant to end-of-life care. There are a number of difficult methodologic issues involved in defining health care markets. Some have strongly advocated the use of the Hospital Referral Regions (HRR),⁹⁸ others the use of network-based measures,^{93,99,100} and others counties. In this project, we have used counties to approximate markets, as has been done in numerous other studies.^{69,89–97,101,102} This was done for a number of reasons: (1) our particular focus is in general hospice use rather than other types of health care utilization; (2) our experience with hospice providers suggests that counties best approximate the way they think about their market's boundaries; (3) empirical tractability and availability of data; and (4) past work suggesting that results are often (but not always) insensitive to the difference between HRRs and counties.¹⁰³

ILLUSTRATION OF THE TYPES OF ANALYSES POSSIBLE

Having assembled the data set, we can now ask some basic questions. In the accompanying paper,¹⁸ we address questions such as the following: Of those in the 50 states and the District of Columbia, newly diagnosed with a serious illness in 1993, how many will have died by the end of 1997? Of those who died, how many used hospice? We are currently engaged in several additional projects based on this data. The questions we are asking include:

- Does the structure of the health care market influence hospice use? This analysis combines the multiple levels of analysis in our data with the power of hierarchical linear models¹⁰⁴ to analyze county-level effects validly while taking into account individual-level differences.
- Does hospice use alter length of life? In order to answer such a question, one must control for the fact that only people close to death are referred to hospice. We exploit the distance to hospice

and distance to hospital measures available in COSI as instrumental variables to study the impact of hospice on duration of life.¹⁰⁵ Similar approaches have been used to evaluate the effectiveness of other medical outcomes.¹⁰⁶

- How does the death of a spouse affect mortality? Because our data contain information about both husbands and wives, we can see how the death of one affects the mortality of the other. Furthermore, we take advantage of differences across diseases in how sensitive probands are to the death of a spouse to try to understand the process by which spouses affect mortality.¹⁰⁷
- Does health care provided to one member of a family affect the health of other members? In this study, we compare the surviving spouses of patients who died using hospice to the spouses of patients who died without hospice. After using a propensity score strategy to address selection issues,¹⁰⁸ we look for differences in mortality of surviving spouses. Such differences have important implications both for meeting patients' wishes and for the ways in which cost effectiveness analysis is conducted.

Certainly other projects are also possible. We hope that COSI offers an opportunity to address a number of questions relevant to the care of our patients—and offers answers that let us improve that care and our social policy.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institute on Aging (R-01 AG15326-01) to N.A.C. and a National Research Service Award from the NIH/National Institute on Aging (T32-AG00243) to T.J.I. We thank Kim Thomas for her expert secretarial assistance and Fay Booker for her contributions to the data analysis presented in this paper.

REFERENCES

1. U.S. Census Bureau: *Statistical Abstract of the United States*. Washington, D.C.: Government Printing Office, 1999.
2. Iwashyna TJ, Chang VW: Racial and ethnic differences in place of death: United States, 1993. *J Amer Geriatric Soc* (in press).

3. The SUPPORT Principal Investigators: A controlled trial to improve care for seriously ill hospitalized patients. *JAMA* 1995;274:1591-1598.
4. Covinsky KE, Goldman L, Cook EF, Oye R, Desbiens N, Reding D, et al: The Impact of Serious Illness on Patient's Families. *JAMA* 1994;272:1839-1844.
5. Riley GF, Lubitz JD: Longitudinal patterns of Medicare use by cause of death. *Health Care Financing Rev* 1989;11:1-12.
6. Riley GF, Potosky AL, Lubitz JD, Kessler LG: Medicare payments from diagnosis to death for elderly cancer patients by stage at diagnosis. *Med Care* 1995;33:828-841.
7. Hogan C, Lunney J, Gabel J, Lynn J: Medicare beneficiaries' costs of care in the last year of life. *Health Affairs* 2001;20:188-195.
8. Weisfeld V, Miller D, Gibson R, Schroeder SA: Improving care at the end of life: What does it take? *Health Affairs* 2000;19:277-283.
9. United States Senate Select Committee on Aging: Barriers to Hospice Care: Are We Shortchanging Dying Patients. Washington, D.C.; 2000.
10. Moyers B: On Our Own Terms: Moyers on Dying. Available at: <www.thirteen.org/onourown/terms>.
11. Christakis NA, Escarce JJ: Survival of Medicare patients following enrollment in hospice programs. *N Engl J Med* 1996;335:172-178.
12. Christakis NA, Iwashyna TJ: The impact of individual and market factors on the timing of initiation of hospice terminal care. *Med Care* 2000;38:528-541.
13. Virnig BA, Kind S, McBean M, Fisher ES: Geographic variation in hospice use prior to death. *J Am Geriatr Soc* 2000;48:1117-1125.
14. Kovar MG, Fitti JE, Chyba MM: The longitudinal study on aging: 1984-90. *Vital Health Statistics* 1992;1(28).
15. Burkhauser RV, Gertler PJ, (eds): *Journal of Human Resources Special Issue: The Health and Retirement Survey, Data Quality and Early Results*. Madison, Wisconsin: University of Wisconsin Press; 1995.
16. Brown C, Duncan GJ, Stafford F: Data watch: The Panel Study of Income Dynamics. *J Econ Perspect* 1996;10:155-168.
17. Adler GS: A profile of the Medicare current beneficiary survey. *Health Care Financing Rev* 1994;15:153-163.
18. Iwashyna TJ, Zhang JX, Christakis NA: Disease-specific patterns of hospice and related healthcare use in an incidence cohort of seriously ill elderly patients. *J Palliat Med* 2002;5:531-538.
19. Hatten J: Medicare's common denominator: The covered population. *Health Care Financing Rev* 1980 (Fall):53-64.
20. McBean AM, Warren JL, Babish JD: Measuring the incidence of cancer in elderly americans using Medicare claims data. *Cancer* 1994;73:2417-2425.
21. Erkinjuntti T, Ostbye T, Steenhuis R, Hachinski V: The effect of different diagnostic criteria on the prevalence of dementia. *N Engl J Med* 1997;337:1667-1674.
22. Krumholz HM, Radford MJ, Wang Y, Chen J, Heiat A, Marciniak TA: National use and effectiveness of the beta-blockers for treatment of elderly patients after acute myocardial infarction. *JAMA* 1998;280:623-629.
23. Iwashyna TJ, Zhang JX, Lauderdale DS, Christakis NA: A methodology for identifying married couples in the Medicare claims data: Mortality, morbidity and health care utilization among the elderly. *Demography* 1998;35:413-419.
24. Iwashyna TJ, Brennan G, Zhang JX, Christakis NA: Finding married couples in the Medicare claims data. *Health Services Outcomes Res Methodol* (in press).
25. Kestenbaum B: Commentary on "A methodology for identifying married couples in the medicare claims data." *Demography* 2000;37:247-249.
26. Iwashyna TJ, Zhang JX, Lauderdale DS, Christakis NA: Refinements of a methodology for detecting married couples in the Medicare claims. *Demography* 2000;37:250-251.
27. Parente ST, Weiner JP, Garnick DW, Richards TM, Fowles J, Lawthers AG, et al: Developing a quality improvement database using health insurance data: A guided tour with application to Medicare's national claims history file. *Am J Med Qual* 1995;10:162-176.
28. Kestenbaum B: A description of the extreme aged population based on improved medicare enrollment data. *Demography* 1992;29:565-580.
29. Korenman S, Goldman N, Fu H: Misclassification bias in estimates of bereavement effects. *Am J Epidemiol* 1997;145:995-1002.
30. Weaver DA: The accuracy of survey-reported marital status: Evidence from survey records matched to Social Security records. *Demography* 2000;37:395-399.
31. U.S. Bureau of the Census: *Current Population Reports, Special Studies, P23-190, 65+ in the United States*. Washington, D.C.: U.S. Government Printing Office, 1996.
32. Clarke SC: Advance report of final marriage statistics, 1989 and 1990. *Monthly Vital Stat Rep* 1995;43(12).
33. Bumpass LL, Sweet JA: National estimates of cohabitation. *Demography* 1989;26:615-626.
34. Chevan A. As Cheaply as One: Cohabitation in the Older Population. *Journal of Marriage and the Family* 1996;58:656-667.
35. Lauderdale DS, Goldberg J: The expanded racial and ethnic codes in the Medicare data files: Their completeness of coverage and accuracy. *Am J Public Health* 1996;86:712-716.
36. Arday SL, Arday DR, Monroe S, Zhang J: HCFA's racial and ethnic data: Current accuracy and recent improvements. *Health Care Financing Rev* 2000;21:107-116.
37. Word DL, Perkins RC Jr: Building a Spanish surname list for the 1990s—A new approach to an old problem. Washington, D.C.: Population Division, U.S. Bureau of the Census; 1996 March 1996. Report No.: Technical Working Paper No. 13.

38. Lauderdale DS, Kestenbaum B: Asian American Ethnic Identification by Surname. *Population Res Policy Rev* 2000;19:283-300.
39. Sandefur GD, Martin M, Eggerling-Boeck J, Mannon SE, Meier AM: An Overview of Racial and Ethnic Demographic Trends. In: Smessler NJ, Wilson WJ, Mitchell F, (eds): *America Becoming: Racial Trends and Their Consequences*. Washington, D.C.: National Academy Press, 2001, pp. 40-102.
40. Charlson ME, Pompei P, Ales KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chron Dis* 1987;40:373-383.
41. Deyo RA, Cherkin DC, Ciol MA: Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45: 613-619.
42. Romano PS, Roos LL, Jollis JG: Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: Differing perspectives. *J Clin Epidemiol* 1993;46:1075-1079.
43. D'Hoore W, Sicotte C, Tilquin C: Risk-adjustment in outcome assessment: The Charlson Comorbidity Index. *Methods Inf Med* 1993;32:382-387.
44. Iezzoni LI, (ed): *Risk Adjustment for Measuring Health Outcomes*, 2 ed. Chicago: Health Administration Press, 1997.
45. Elixhauser A, Steiner C, Harris DR, Coffey RM: Comorbidity measures for use with administrative data. *Med Care* 1998;36:8-27.
46. Schwartz M, Iezzoni LI, Moskowitz MA, Ash AS, Sawitz E: The importance of comorbidities in explaining differences in patient costs. *Med Care* 1996;34:767-782.
47. Iezzoni LI, Daley J, Heeren T, Filey SM, Hughes JS, Fisher ES, et al: Using administrative data to screen hospitals for high complication rates. *Inquiry* 1994;31:40-55.
48. Kuykendall DH, Ashton CM, Johnson ML, Geraci JM: Identifying complications and low provider adherence to normative practice using administrative data. *Health Serv Res* 1995;30:531-554.
49. DesHarnais SI, McMahon LF Jr., Wroblewski RT, Hogan AJ: Measuring hospital performance: The development and validation of risk-adjusted indexes of mortality, readmissions, and complications. *Med Care* 1990;28:1127-1141.
50. Brailer DJ, Kroch E, Pauly MV, Huang J: Comorbidity-adjusted complication risk. *Med Care* 1996;34: 490-505.
51. Weiner JP, Starfield BH, Steinwachs DM, Mumford LM: Development and application of a population-oriented measure of ambulatory care case-mix. *Med Care* 1991;29:452-472.
52. Starfield B, Weiner J, Mumford L, Steinwachs D: Ambulatory care groups: A categorization of diagnoses for research and management. *Health Serv Res* 1991;26:53-74.
53. Fowles JB, Weiner JP, Knutson D, Fowler E, Tucker AM, Ireland M: Taking health status into account when setting capitation rates: A comparison of risk-adjustment methods. *JAMA* 1996;276:1316-1321.
54. D'Hoore W, Bouckaert A, Tilquin C: Practical considerations on the use of the Charlson Comorbidity Index with administrative data bases. *J Clin Epidemiol* 1996;49:1429-1433.
55. Roos NP, Wennberg JE, Malenka DJ, Fisher ES, McPherson K, Anderson TF, et al: Mortality and reoperation after open and transurethral resection of the prostate for benign prostatic hyperplasia. *N Engl J Med* 1989;320:1120-1124.
56. Romano PS, Roos LL, Jollis JG: Further evidence concerning the use of a clinical comorbidity index with ICD-9-CM administrative data. *J Clin Epidemiol* 1993;46:1085-1090.
57. Roos LL, Sharp SM, Cohen MM: Comparing clinical information with claims data: Some similarities and differences. *J Clin Epidemiol* 1991;44:881-888.
58. Hughes JS, Iezzoni LI, Daley J, Greenberg L: How severity measures rate hospitalized patients. *J Gen Intern Med* 1996;11:303-311.
59. Ghali WA, Hall RE, Rosen AK, Ash AS, Moskowitz MA: Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data. *J Clin Epidemiol* 1996;49:273-278.
60. Zhang JX, Iwashyna TJ, Christakis NA: The impact of alternative lookback periods and sources of information on Charlson comorbidity adjustment in Medicare claims. *Med Care* 1999;37:1128-1139.
61. American Hospital Association: *The Annual Survey of Hospitals Database Documentation for 1993 Data*. Chicago: American Hospital Association, 1994.
62. Hill CA, Rudolph BA: *The Index of Hospital Quality: 1993 Survey on Best Hospitals*. Chicago: NORC, 1993.
63. Hill CA, Winfrey KL, Rudolph BA: "Best hospitals": A description of the methodology for the index of hospital quality. *Inquiry* 1997;34:80-90.
64. Chen J, Radford MJ, Wang Y, Marciniak TA, Krumholz HM: Do "America's best hospitals" perform better for acute myocardial infarction? *N Engl J Med* 1999;340:286-292.
65. Escarce J, Epstein K, Colby D, Schwartz J: Racial differences in the elderly's use of medical procedures and diagnostic tests. *Am J Public Health* 1993;83: 948-954.
66. Clark WD, Hulbert MM: Research issues: Dually eligible Medicare and Medicaid beneficiaries, challenges and opportunities. *Health Care Financing Rev* 1998;20:1-10.
67. Philbin EF, McCullough PA, DiSalvo TG, Dec GW, Jenkins PL, Weaver WD: Socioeconomic status is an important determinant of the use of invasive procedures after acute myocardial infarction in New York State. *Circulation (online)* 2000;102(19 Suppl 3): III107-15.
68. U.S. Department of Commerce Bureau of the Census. *Census of Population and Housing, 1990, United States: Summary Tape File 3B*. In: U.S. Department of Commerce Bureau of the Census, 1991.
69. Roetzheim RG, Pal N, Gonzalez EC, Ferrante JM,

- Van Durme DJ, Ayanian JZ, et al: The effects of physician supply on the early detection of colorectal cancer. *J Family Pract* 1999;48:850-858.
70. Carlisle DM, Leake BD: Differences in the effect of patients' socioeconomic status on the use of invasive cardiovascular procedures across health insurance categories. *Am J Public Health* 1998;88:1089-1092.
 71. Alexander GC, Sehgal AR: Barriers to cadaveric renal transplantation among blacks, women, and the poor. *JAMA* 1998;280:1148-1152.
 72. Feinglass J, Kaushik S, Handel D, Kosifas A, Martin GJ, Pearce WH: Peripheral bypass surgery and amputation: Northern Illinois demographics, 1993 to 1997. *Arch Surg* 2000;135:75-80.
 73. Kaestner R, Racine A, Joyce T: Did recent expansions in Medicaid narrow socioeconomic differences in hospitalization rates of infants? *Med Care* 2000;38:195-206.
 74. Garcia JA, Romano PS, Chan BK, Kass PH, Robbins JA: Sociodemographic factors and the assignment of do-not-resuscitate orders in patients with acute myocardial infarctions. *Med Care* 2000;38:670-678.
 75. Hofer TP, Wolfe RA, Tedeschi PJ, MacMahon LF, Griffith JR: Use of community versus individual socioeconomic data in predicting variation in hospital Use. *Health Serv Res* 1998;33:243-259.
 76. Krieger N: Overcoming the absence of socioeconomic data in medical records: Validation and application of a census-based methodology. *Am J Public Health* 1992;82:703-710.
 77. Geronimus AT, Bound J, Neidert LJ: On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics. *J Am Statistical Assoc* 1996;91:529-537.
 78. Geronimus AT, Bound J: Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. *Am J Epidemiol* 1998;148:475-486.
 79. Robinson WS: Ecological correlations and the behavior of individuals. *Am Sociol Rev* 1950;15:351-357.
 80. Davey Smith G, Ben-Shlomo Y, Hart C: Re: Use of Census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. *Am J Epidemiol* 1999;150:996-999.
 81. Greenwald HP, Polissar NL, Borgatta EF, McCorkle R: Detecting survival effects of socioeconomic status: Problems in the use of aggregate measures. *J Clin Epidemiol* 1994;47:903-909.
 82. Hyndman JC, Holman CD, Hockey RL, Donovan RJ, Corti B, Rivera J: Misclassification of social disadvantage based on geographical areas: comparison of postcode and collector's district analyses. *Int J Epidemiol* 1995;24:165-176.
 83. Summer AA, Wolfe BL: Estimating household income from location. *J Am Statistical Assoc* 1978;73:288-292.
 84. Krieger N, Gordon D: Re: Use of Census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. *Am J Epidemiol* 1999;150:892-896.
 85. Oliver ML, Shapiro TM: *Black Wealth/White Wealth: A New Perspective on Racial Inequality*. New York: Routledge, 1995.
 86. Wilson WJ: *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press, 1987.
 87. Massey DS: The age of extremes: Concentrated affluence and poverty in the twenty-first century. *Demography* 1996;33:395-412.
 88. Berkman LF, Kawachi I, (eds): *Social Epidemiology*. Oxford: Oxford University Press, 2000.
 89. Kerstein J, Pauly MV, Hillman A: Primary care physician turnover in HMOs. *Health Services Res* 1994;29:17-37.
 90. Banaszak-Holl J, Zinn JS, Mor V: The impact of market and organizational characteristics on nursing care facility service innovation: A resource dependency perspective. *Health Services Res* 1996;31:97-117.
 91. Halfon N, Newacheck PW, Wood DL, St Peter RF: Routine emergency department use for sick care by children in the United States. *Pediatrics* 1996;98:28-34.
 92. Hartley D, Moscovice I, Christianson J: Mobile technology in rural hospitals: The case of the CT scanner. *Health Services Res* 1996;31:213-234.
 93. Succi MJ, Lee SY, Alexander JA: Effects of market position and competition on rural hospital closures. *Health Services Res* 1997;31:679-699.
 94. Wholey DR, Christianson JB, Engberg J, Bryce C: HMO market structure and performance: 1985-1995. *Health Affairs* 1997;16:75-84.
 95. Lafata JE, Koch GG, Weissert WG: Estimating activity limitation in the noninstitutionalized population: A method for small areas. *Am J Public Health* 1994;84:1813-1817.
 96. Lambrew JM, Ricketts TC: Patterns of obstetrical care in single-hospital, rural counties. *Med Care* 1993;31:822-833.
 97. Mullan F, Politzer RM, Davis CH: Medical migration and the physician workforce. *International medical graduates and American medicine*. *JAMA* 1995;273:1521-1527.
 98. Wennberg JE, Cooper MM, (eds). *The Dartmouth Atlas of Health Care*. American Hospital Publishing, Chicago, IL: 1998.
 99. Phibbs CS, Robinson JC: A variable-radius measure of local hospital market structure. *Health Services Res* 1993;28:313-324.
 100. Sohn M-W: *From Regional to Local Markets: Network Study of Competition in California Hospital Markets*. Chicago: University of Chicago, 1996.
 101. Murtaugh CM: Discharge planning in nursing homes. *Health Services Res* 1994;28:751-769.
 102. Padgett DK, Patrick C, Burns BJ, Schlesinger HJ: Ethnic differences in use of inpatient mental health services by blacks, whites, and Hispanics in a national

- insured population. *Health Services Res* 1994;29:135-153.
103. McLaughlin CG, Normolle DP, Wolfe RA, McMahon LF Jr., Griffith JR: Small-area variation in hospital discharge rates: Do socioeconomic variables matter? *Med Care* 1989;27:507-521.
 104. Bryk AS, Raudenbush SW: *Hierarchical Linear Models*. Newbury Park, CA: Sage 1992.
 105. Harvey DJ: *Instrumental Variables in Survival Analysis*. In: University of Chicago Dissertation Library. Chicago, IL: 2001.
 106. McClellan M, McNeil BJ, Newhouse JP: Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA* 1994;272:859-866.
 107. Iwashyna TJ: *In Sickness and In Health: Understanding the Effects of Marriage on Health*. In: University of Chicago Dissertation Library. Chicago, IL: 2001.
 108. Rosenbaum PR, Rubin DB: The central role of the propensity score in observational studies for causal effects. *Biometrika* 1984;70:41-55.
 109. Jollis JG, DeLong ER, Peterson ED, Muhlbaier LH, Fortin DF, Califf RM, et al: Outcome of acute myocardial infarction according to the specialty of the admitting physician. *N Engl J Med* 1996;335:1880-1887.
 110. Asch S, Sloss E, Kravitz R, Kamberg C, Genovese B, Young R: Access to care for the elderly Project. Santa Monica: RAND; 1995 July 1995. Report No.: PM-435-PPRC.
 111. Draper D, Kahn KL, Reinisch EJ, Sherwood MJ, Carney MF, Kosecoff J, et al: Studying the effects of the DRG-based prospective payment system on quality of care: Design, sampling and fieldwork. *JAMA* 1990;264:1956-1961.
 112. Fisher ES, Whaley FS, Krushat M, Malenka DJ, Fleming C, J.A. B, et al. The accuracy of Medicare's hospital claims data: Progress has been made, but problems remain. *Am J Public Health* 1992;82:243-248.
 113. Taylor DH Jr., Whellan DJ, Sloan FA: Effects of admission to a teaching hospital on the cost and quality of care for Medicare beneficiaries. *N Engl J Med* 1999;340:293-299.
 114. Krumholz HM, Parent EM, Tu N, Vaccarino V, Wang Y, Radford MJ, et al: Readmission after Hospitalization for Congestive Heart Failure among Medicare Beneficiaries. *Arch Intern Med* 1997;157:99-104.
 115. Lauderdale DS, Jacobsen SJ, Furner SE, Levy PS, Brody JA, Goldberg J: Hip fracture incidence among elderly Asian-American populations. *Am J Epidemiol* 1997;146:502-509.
 116. Benesch C, Witter DM Jr., Wilder AL, Duncan PW, Samsa GP, Matchar DB: Inaccuracy of the International Classification of Disease (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. *Neurology* 1997;49:660-664.
 117. Holloway RG, Witter Jr. DM, Lawton KB, Lipscomb J, Samsa G: Inpatient costs of specific cerebrovascular events at five academic medical centers. *Neurology* 1996;46:854-860.
 118. Wolinsky FD, Wan GJ, Gurney JG, Bentley DW: The risk of hospitalization for ischemic stroke among older adults. *Med Care* 1998;36:449-461.
 119. Lee AJ, Huber J, Stason WB: Poststroke rehabilitation in older Americans: The Medicare experience. *Med Care* 1996;34:811-825.
 120. McBean AM, Babish JD, Warren JL: Determination of lung cancer incidence in the elderly using Medicare claims data. *Am J Epidemiol* 1993;137:226-234.
 121. Ries LAG, Kosary CL, Hankey BF, Miller BA, Hiras A, Edwards BK, (eds): *SEER Cancer Statistics Review, 1973-1994*. Bethesda, MD: National Cancer Institute. NIH Pub. No. 97-2789; 1997.
 122. Counsell CE, Collie DA, Grant R: Limitations of using a cancer registry to identify incident primary intracranial tumors. *J Neurol, Neurosurg, Psychiatry* 1997;63:94-97.
 123. Loomis DP, Savitz DA: Mortality from brain cancer and leukaemia among electrical workers. *Br J Industrial Med* 1990;47:633-638.
 124. McKinney PA, Ironside JW, Harkness EF, Arango JC, Doyle D, Black RJ: Registration quality and descriptive epidemiology of childhood brain tumours in Scotland 1975-90. *Br J Cancer* 1994;70:973-979.
 125. Allison P, Locker D, Feine JS: The role of diagnostic delays in the prognosis of oral cancer: A review of the literature. *Oral Oncol* 1998;34:161-170.
 126. Allison P, Franco E, Feine J: Predictors of professional diagnostic delays for upper aerodigestive tract carcinoma. *Oral Oncol* 1998;34:127-132.
 127. Ventrees J, Manton KG: The complexity of chronic disease at later stages: Practical implications for prospective payment and data collection. *Inquiry* 1986;23:154-165.

Address reprint requests to:

Nicholas A. Christakis, M.D., Ph.D., MPH
 Department of Health Care Policy
 Harvard Medical School
 180 Longwood Avenue
 Boston, MA 02115

E-mail: christakis@hcp.med.harvard.edu

This article has been cited by:

1. Amber E. Barnato, Elan D. Cohen, Keili A. Mistovich, Chung-Chou H. Chang. 2015. Hospital End-of-Life Treatment Intensity Among Cancer and Non-Cancer Cohorts. *Journal of Pain and Symptom Management* **49**, 521-529.e5. [[CrossRef](#)]
2. Woolley Douglas C., Old Jerry L., Zackula Rosalee E., Davis Nancy. 2013. Acute Hospital Admissions of Hospice Patients. *Journal of Palliative Medicine* **16**:12, 1515-1522. [[Abstract](#)] [[Full Text HTML](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)] [[Supplemental Material](#)]
3. Stephen R. Connor, Bruce Pyenson, Kathryn Fitch, Carol Spence, Kosuke Iwasaki. 2007. Comparing Hospice and Nonhospice Patient Survival Among Patients Who Die Within a Three-Year Window. *Journal of Pain and Symptom Management* **33**, 238-246. [[CrossRef](#)]
4. Ming Wen, Nicholas A. Christakis. 2006. Prospective effect of community distress and subcultural orientation on mortality following life-threatening diseases in later life. *Sociology of Health and Illness* **28**:10.1111/shil.2006.28.issue-5, 558-582. [[CrossRef](#)]
5. M WEN, K CAGNEY, N CHRISTAKIS. 2005. Effect of specific aspects of community social environment on the mortality of Individuals diagnosed with serious illness. *Social Science & Medicine* **61**, 1119-1134. [[CrossRef](#)]
6. Ming Wen, Nicholas A. Christakis. 2005. Neighborhood Effects on Posthospitalization Mortality: A Population-based Cohort Study of the Elderly in Chicago. *Health Services Research* **40**:10.1111/hesr.2005.40.issue-4, 1108-1127. [[CrossRef](#)]
7. Theodore J. Iwashyna. 2004. Critical Care Use during the Course of Serious Illness. *American Journal of Respiratory and Critical Care Medicine* **170**, 981-986. [[CrossRef](#)]
8. B PYENSON, S CONNOR, K FITCH, B KINZBRUNNER. 2004. Medicare cost in matched hospice and non-hospice cohorts. *Journal of Pain and Symptom Management* **28**, 200-210. [[CrossRef](#)]
9. T Iwashyna. 2003. Marriage, widowhood, and health-care use. *Social Science & Medicine* **57**, 2137-2147. [[CrossRef](#)]
10. N Christakis. 2003. The health impact of health care on families: a matched cohort study of hospice use by decedents and mortality outcomes in surviving, widowed spouses. *Social Science & Medicine* **57**, 465-475. [[CrossRef](#)]
11. Theodore J Iwashyna, Virginia W Chang, James X Zhang, Nicholas A Christakis. 2002. The Lack of Effect of Market Structure on Hospice Use. *Health Services Research* **37**:10.1111/hesr.2002.37.issue-6, 1531-1551. [[CrossRef](#)]
12. Theodore J. Iwashyna, James X. Zhang, Nicholas A. Christakis. 2002. Disease-Specific Patterns of Hospice and Related Healthcare Use in an Incidence Cohort of Seriously Ill Elderly Patients. *Journal of Palliative Medicine* **5**:4, 531-538. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]