MEDICAL CARE
Volume 36, Number 1, pp 8–27

Comorbidity Measures for Use with Administrative Data

ANNE ELIXHAUSER, PHD,* CLAUDIA STEINER, MD, MPH,† D. ROBERT HARRIS, PHD,‡ AND ROSANNA M. COFFEY, PHD,§

OBJECTIVES. This study attempts to develop a comprehensive set of comorbidity measures for use with large administrative inpatient datasets.

METHODS. The study involved clinical and empirical review of comorbidity measures, development of a framework that attempts to segregate comorbidities from other aspects of the patient's condition, development of a comorbidity algorithm, and testing on heterogeneous and homogeneous patient groups. Data were drawn from all adult, nonmaternal inpatients from 438 acute care hospitals in California in 1992 (n = 1,779,167). Outcome measures were those commonly available in administrative data: length of stay, hospital charges, and in-hospital death.

RESULTS. A comprehensive set of 30 comorbidity measures was developed. The comorbidities were associated with substantial increases in length of stay, hospital charges, and mortality both for heterogeneous and homogeneous disease groups. Several comorbidities are described that are important predictors of outcomes, yet commonly are not measured. These include mental disorders, drug and alcohol abuse, obesity, coagulopathy, weight loss, and fluid and electrolyte disorders.

CONCLUSIONS. The comorbidities had independent effects on outcomes and probably should not be simplified as an index because they affect outcomes differently among different patient groups. The present method addresses some of the limitations of previous measures. It is based on a comprehensive approach to identifying comorbidities and separates them from the primary reason for hospitalization, resulting in an expanded set of comorbidities that easily is applied without further refinement to administrative data for a wide range of diseases.

Key words: comorbidity; administrative data; hospital resources; in-hospital mortality. (Med Care 1998;36:8–27)

Measures of the overall medical condition of patients are essential for health care research, whether collecting data prospectively or using data that have been collected for another purpose. This is true for testing new treatments, assessing established ones, evaluating health plans and providers, or studying the impact of health care policies. Biomedical evaluations that have used randomized controlled trials usually have excluded patients with certain preexisting conditions, although even experimental research increasingly has been using statistical controls on more heterogeneous populations.[1] Outcomes assessments of clinical procedures applied to large populations of patients typically have used statistical techniques to control retrospectively for clinical differences among patients.[2–4] Comparisons of health care providers have attempted to adjust for the medical and financial risk of serving different populations.[5] Health policy studies, such

*From MEDTAP International, Inc., Bethesda, Maryland.

†From the Agency for Health Care Policy and Research Center for Organization & Delivery Systems, Rockville, Maryland.

‡From WESTAT, Inc., Rockville, Maryland.

§From The MEDSTAT Group, Inc., Washington, DC.

This work was conducted while the authors were employees of the Agency for Health Care Policy and Research (AHCPR). The views expressed in this article are those of the authors and do not necessarily reflect those of AHCPR.

Additional materials are available from the authors.

Address correspondence to: Anne Elixhauser, PhD, MEDTAP International, Inc., 7101 Wisconsin Ave., Suite 600, Bethesda, MD 20814; e-mail: elix@medtap.com.

Received April 10, 1997; initial review completed May 20, 1997; final acceptance July 29, 1997.

as evaluations of the effects of payment policies or assessments of the performance of plans and providers, have used statistical techniques to control for the medical conditions of the heterogeneous patient populations that inevitably must be compared.[6,7]

When using administrative data, preexisting conditions, or comorbidities, should always be controlled. In general, comorbid conditions have been handled analytically by: (1) stratifying patients into groups—those with a comorbidity and those without; (2) using separate binary indicators for discrete conditions; or (3) summarizing comorbidity information into an index or score that provides a single parameter for measuring multiple comorbidities.[1,8–15]

One of the most commonly used indexes was developed by Charlson et al.[1] Although it was developed for the express purpose of prospectively predicting 1 year mortality among patients being considered for breast cancer clinical trials, it has been applied to discharge abstract and claims data to predict short-term outcomes such as in-hospital mortality, blood transfusions, hospitalization charges, and length of stay.[16–20]

Romano et al[21,22] explored the Charlson Index and pointed out a number of precautions when using indexes to control for comorbidities. First, the complexity of ICD-9-CM coding and coding idiosyncrasies must be taken into account in defining comorbidities.[23] Second, the weights for particular comorbidities should be estimated separately for different populations and different outcomes because their predictive values differ by patient groups. Finally, there is no evidence that the comorbidities included in the Charlson Index are comprehensive, given that it includes only those conditions that happened to occur in a narrowly defined clinical population of fewer than 600 patients.

Purpose

This project attempts to improve on measures of comorbidity for use with administrative inpatient databases. This study used a large administrative data set to develop and test comorbidity measures that can be used to control for a broad array of patients' underlying, preexisting conditions in many types of studies. Because research using administrative datasets often examines resource use or clinical outcomes, we developed the comorbidity measures to predict hospital charges, length of stay, and in-hospital mortality.

Methods

Defining Important Comorbidities

The entire burden of illness for a patient, as reflected in the information relevant to a hospitalization, can be divided into five separate concepts:

1. The primary reason for hospitalization, as reflected in the principal diagnosis;
2. The severity of the principal diagnosis;
3. Complications that result from the process of care;
4. Unimportant comorbidities or other conditions present on admission that have a trivial impact on resource use and outcomes; and
5. Important comorbidities or conditions present on admission that are not related directly to the main reason for hospitalization, but that increase the intensity of resources used or increase the likelihood of a poor outcome.

To conceptually identify important comorbidities, we attempted to exclude information that relates to the other aspects of a patient's condition, concepts 1 through 4 above. We strove to eliminate the main reason for hospitalization by restricting our search for comorbidities to secondary diagnoses. We considered a secondary diagnosis to be a comorbidity only when it was not directly related to the principal diagnosis. To apply this restriction, we considered secondary diagnoses to be comorbidities only when they did not relate directly to the diagnosis-related group (DRG) assignment of each patient.[24] For example, if the ICD-9-CM code 428.9 (heart failure, unspecified) appeared on the record, it was counted as a comorbidity only if the record did not fall into any of the cardiac DRGs. The secondary diagnosis of heart failure under a cardiac DRG most likely represents a further specification of the principal diagnosis and is not likely to be a discrete and separate coexisting condition. By implementing this rule, we assumed that heart failure under a cardiac DRG is a modifier of the principal diagnosis and an indicator of illness severity that can be taken into account by using severity of illness measures such as Disease Staging.[25] Although this assumption may not always be accurate, we chose to take a more conservative approach to identifying comorbidities that did not overlap

with other established and often used approaches to assessing disease severity.

We attempted to eliminate complications by excluding ICD-9-CM codes that reflect acute conditions that could result from medical misadventures. A number of diagnoses that might be considered comorbidities were not counted as such because the conditions were not distinguishable from complications that might have originated during the hospitalization as a result of diagnostic or therapeutic interventions. Such acute conditions include pneumonia, pleural effusion, urinary tract infection, cardiac arrest, cardiogenic shock, and respiratory failure. Again, this represented a conservative approach in that these conditions could be comorbidities, but given their acute nature, they are often complications.

Finally, we eliminated unimportant comorbidities because it has been demonstrated in the literature and through empirical work reported here that these conditions do not have a significant impact on resource use or mortality if they are not the principal diagnosis. Such conditions include benign prostatic hypertrophy, inguinal hernia, and diverticulosis.[12,26]

We defined comorbidity as a clinical condition that exists before a patient's admission to the hospital, is not related to the principal reason for the hospitalization, and is likely to be a significant factor influencing mortality and resource use in the hospital. Although the first part of this definition cannot be strictly met using most discharge abstract data, the other two parts can be addressed to some extent.

Selecting and Refining the Comorbidity List

An initial list of 41 comorbidities was selected through review of published studies identifying comorbid conditions and by examining the ICD-9-CM coding manual for additional comorbid conditions that met our definition of a comorbidity.[1,19,27–31] To refine the initial list of comorbidities, a series of univariate and multivariate analyses were conducted. Comorbidities that were infrequent (late effects of infectious illness, other heart disease) or statistically unrelated to length of stay, total charges, or in-hospital mortality (osteoarthritis, old myocardial infarction, cardiomyopathy, cerebrovascular disease, dementia, other endocrine disorders, renal disease, leukemia, other anemias, inflammatory bowel disease, atherosclerosis) were excluded. Comorbidities that were too heterogeneous were partitioned further (anemia was split into blood loss and deficiency anemias; hypothyroidism was split from other endocrine disorders; alcohol and drug abuse were split into separate comorbidities, and mental illness was split into psychoses and depression). Similar comorbidities that, when separate, were weakly related to the outcome measures, but, when merged, had a stronger relationship were combined (hypertension and complicated hypertension, mild and severe liver disease). Finally, an algorithm was created to avoid double-counting closely related comorbidities. For example, if a patient's record included both diabetes and diabetes complications, only the more severe comorbidity (diabetes complications) was counted. The above revisions resulted in a final set of 30 comorbidities. The final list of comorbidities, including their ICD-9-CM codes, the DRGs that screen out comorbidities directly related to the reason for hospitalization, and the rules for eliminating double-counting are displayed in Table 1.

Study Population

The study population was drawn from all inpatient hospital stays in California for the year 1992. These data were part of the Statewide Inpatient Database (SID) of the Healthcare Cost and Utilization Project, a research project of the Agency for Health Care Policy and Research. California was chosen because it is the most populous state in the nation and because California collects more clinical information than any other state—the principal diagnosis and procedure and up to 29 secondary diagnoses and procedures.

The California SID contains inpatient stay records (N = 3,597,735) from all 439 non-federal, community hospitals in the state. The discharge record includes typical patient-level clinical, demographic, and resource use information. The study population was restricted to nonmaternal records of patients 18 years and older who were admitted to acute care hospitals and were not discharged to a long-term care facility or another hospital (n = 1,779,167). Patients discharged to another institution were excluded because it was assumed that their episode of care extended beyond this hospitalization, thus their use of resources during this hospitalization would have been truncated, and their inclusion would have introduced bias. The study population was drawn from 438 hospitals and is described in Table 2.

Approach

The list of comorbidities was evaluated on the heterogenous set of records with all reasons for hospitalizations and on homogeneous subsets of these hospitalization records for patients with particular diseases. The heterogeneous group allowed us to evaluate the usefulness of the comorbidities as controls for studies that involve a broad spectrum of cases, for example, studies of all discharges from a particular hospital.

The homogeneous groups allowed us to test the usefulness of the comorbidities for studies of specific clinical populations. We selected 10 conditions representing acute, chronic, surgical, or nonsurgical conditions: breast cancer ($n = 15,968$), acute myocardial infarction ($n = 39,686$), asthma ($n = 21,895$), appendicitis ($n = 20,438$), abdominal hernia ($n = 12,746$), diverticulosis and diverticulitis ($n = 14,064$), biliary tract disease ($n = 53,551$), low back pain ($n = 45,467$), pneumonia ($n = 70,333$), and diabetes mellitus with complications ($n = 26,287$).

Ordinary least square regressions were used to assess the contribution of the comorbidities to predicting logged length of stay and logged total charges. Length of stay and charges were transformed into logarithms because they had skewed distributions with very long tails of few cases with lengthy or expensive hospitalizations. Logistic regression was used to assess the contribution of the comorbidities to predicting in-hospital mortality. Each regression model contained a set of independent variables that are commonly used to isolate the influence of demographics, financial incentives, and clinical differences among patients, including: age (in years), race (black, other, white = reference category), gender (female = reference), expected primary payer (Medicare, Medicaid, self-pay, other, private insurance = reference), emergency admission (urgent and elective combined = reference), surgery performed in the hospital (ie, patient assigned any surgical DRG), and presence of any complication that occurred in the hospital. Complications were based on diagnoses, defined by DesHarnais,[32] that indicate adverse events resulting from medical treatments or procedures. The ICD-9-CM codes for complications are listed in a footnote to Table 2. In defining complications, we avoided potential comorbidities as defined by the DRG system (eg, MI or anemias), thus we may have underestimated the true rate of complications. Complications never overlapped with our definitions of comorbidities.

Results

Heterogeneous Group

In this population of adult, nonmaternal discharges with varied reasons for hospitalization, nearly 60% of patients had at least one comorbidity. Table 3 shows that the greater the number of comorbidities affecting a patient, the greater the resource use and the likelihood of death in the hospital. Length of stay and total charges were twice as high for patients with three or more comorbidities compared with patients who had no comorbidities, and the death rate was higher by a factor of seven. As expected, the number of comorbidities and the age of the patient also were positively related.

Table 4 reveals that the most frequent comorbidities among all discharges were: hypertension (17.9% of cases), fluid and electrolyte disorders (13.3%), chronic pulmonary disease (9.9%), diabetes (7.8%), deficiency anemias (7.3%), and cardiac arrhythmias (6.8%).

Table 4 also shows the independent effects of each comorbidity on the outcomes of interest, while holding constant the effects of all other comorbidities and other variables specified in the model. The antilogarithm of the coefficient, or the transformed coefficient, is the percentage change in resource use from having a particular comorbidity, independent of other patient characteristics. The odds ratio is the odds that a patient with a particular comorbidity will die in the hospital, holding constant all other factors measured. For example, congestive heart failure increased both length of stay and charges by 35% and increased the likelihood of death by 2.3 times compared with patients who had no congestive heart failure recorded but were similar on the other dimensions measured.

In general, if a comorbidity was associated with a statistically significant increase in length of stay or total hospital charges, it usually increased the odds of dying in the hospital. There were, however, some interesting exceptions to this observation. Although hypertension increased hospital charges by 6%, it decreased the odds of in-hospital death by 40%. This finding is consistent with other studies and is an indicator of a bias in discharge abstract coding—namely, the severity of the patient's condition inversely affected the coding of certain common and singly unthreatening conditions.[6,9] A seriously ill pa-

TABLE 1. Definitions of Comorbidities

Comorbidity	ICD-9-CM Codes	DRG Screen: Case Does *Not* Have the Following Disorders (DRG):
1. Congestive heart failure	398.91, 402.11, 402.91, 404.11, 404.13, 404.91, 404.93, 428.0–428.9	Cardiac[a]
2. Cardiac arrhythmias	426.10, 426.11, 426.13, 426.2–426.53, 426.6–426.89, 427.0, 427.2, 427.31, 427.60, 427.9, 785.0, V45.0, V53.3	Cardiac[a]
3. Valvular disease	093.20–093.24, 394.0–397.1, 424.0–424.91, 746.3–746.6, V42.2, V43.3	Cardiac[a]
4. Pulmonary circulation disorders	416.0–416.9, 417.9	Cardiac[a] or COPD (88)
5. Peripheral vascular disorders	440.0–440.9, 441.2, 441.4, 441.7, 441.9, 443.1–443.9, 447.1, 557.1, 557.9, V43.4	Peripheral vascular (130–131)
6. Hypertension (*combined*)		
Hypertension, uncomplicated	401.1, 401.9	Hypertension (134)
Hypertension, complicated	402.10, 402.90, 404.10, 404.90, 405.11, 405.19, 405.91, 405.99	Hypertension (134) or cardiac[a] or renal[a]
7. Paralysis	342.0–342.12, 342.9–344.9	Cerebrovascular (5, 14–17)
8. Other neurological disorders	331.9, 332.0, 333.4, 333.5, 334.0–335.9, 340, 341.1–341.9, 345.00–345.11, 345.40–345.51, 345.80–345.91, 348.1, 348.3, 780.3, 784.3	Nervous system (1–35)
9. Chronic pulmonary disease	490–492.8, 493.00–493.91, 494, 495.0–505, 506.4	COPD (88) or asthma (96–98)
10. Diabetes, uncomplicated[b]	250.00–250.33	Diabetes (294–295)
11. Diabetes, complicated[b]	250.40–250.73, 250.90–250.93	Diabetes (294–295)
12. Hypothyroidism	243–244.2, 244.8, 244.9	Thyroid (290) or endocrine (300–301)
13. Renal failure	403.11, 403.91, 404.12, 404.92, 585, 586, V42.0, V45.1, V56.0, V56.8	Kidney transplant (302) or renal failure/dialysis (316–317)
14. Liver disease	070.32, 070.33, 070.54, 456.0, 456.1, 456.20, 456.21 571.0, 571.2, 571.3, 571.40–571.49, 571.5, 571.6, 571.8, 571.9, 572.3, 572.8, V42.7	Liver[a]
15. Peptic ulcer disease excluding bleeding	531.70, 531.90, 532.70, 532.90, 533.70, 533.90, 534.70, 534.90, V12.71	GI hemorrhage or ulcer (174–178)
16. AIDS[b]	042–044.9	HIV (488–490)

(Continues)

tient may have had so many medical problems that hypertension, although detected, was not abstracted for the discharge record; at the same time, a relatively healthy patient may have been more likely to have had hypertension recorded if it existed because there were fewer serious clinical findings to record in the abstract and medical record. Similar anomalies were evident for the following comorbidities: valvular heart disease, hypothyroidism, peptic ulcer disease without hemorrhage, obesity, blood loss anemia, and depression. In these cases, the comorbidity was associated with longer lengths of stay and higher charges, but a lower odds of in-hospital mortality.

Table 4 also shows the importance of the comorbidities as a group in predicting resource use. At the end of the table are R^2 results of excluding all comorbidities and then including them as a group in the complete model. In the regressions, the proportion of explained variation in length of

TABLE 1. (Continued)

Comorbidity	ICD-9-CM Codes	DRG Screen: Case Does *Not* Have the Following Disorders (DRG):
17. Lymphoma	200.00–202.38, 202.50–203.01, 203.8–203.81, 238.6, 273.3, V10.71, V10.72, V10.79	Leukemia/lymphoma[a]
18. Metastatic cancer[b]	196.0–199.1	Cancer[a]
19. Solid tumor without metastasis[b]	140.0–172.9, 174.0–175.9, 179–195.8, V10.00–V10.9	Cancer[a]
20. Rheumatoid arthritis/collagen vascular diseases	701.0, 710.0–710.9, 714.0–714.9, 720.0–720.9, 725	Connective tissue (240–241)
21. Coagulopathy	2860–2869, 287.1, 287.3–287.5	Coagulation (397)
22. Obesity	278.0	Obesity procedure (288) or nutrition/metabolic (296–298)
23. Weight loss	260–263.9	Nutrition/metabolic (296–298)
24. Fluid and electrolyte disorders	276.0–276.9	Nutrition/metabolic (296–298)
25. Blood loss anemia	2800	Anemia (395–396)
26. Deficiency anemias	280.1–281.9, 285.9	Anemia (395–396)
27. Alcohol abuse	291.1, 291.2, 291.5, 291.8, 291.9, 303.90–303.93, 305.00–305.03, V113	Alcohol or drug (433–437)
28. Drug abuse	292.0, 292.82–292.89, 292.9, 304.00–304.93, 305.20–305.93	Alcohol or drug (433–437)
29. Psychoses	295.00–298.9, 299.10–299.11	Psychoses (430)
30. Depression	300.4, 301.12, 309.0, 309.1, 311	Depression (426)

ICD-9-CM, International Classification of Diseases, 9th Revision, Clinical Modification; DRG, diagnosis-related group; COPD, chronic obstructive pulmonary disease; GI, gastrointestinal; AIDS, acquired immune deficiency syndrome; HIV, human immunodeficiency virus.

[a]Definitions of DRG groups: Cardiac: DRGs 103–108, 110–112, 115–118, 120–127, 129, 132–133, 135–143; Renal: DRGs 302–305, 315–333; Liver: DRGs 199–202, 205–208; Leukemia/lymphoma: DRGs 400–414, 473, 492; Cancer: DRGs 10, 11, 64, 82, 172, 173, 199, 203, 239, 257–260, 274, 275, 303, 318, 319, 338, 344, 346, 347, 354, 355, 357, 363, 366, 367, 406–414.

[b]A hierarchy was established between the following pairs of comorbidities: If both uncomplicated diabetes and complicated diabetes are present, count only complicated diabetes. If both solid tumor without metastasis and metastatic cancer are present, count only metastatic cancer.

stay (R^2) more than doubled from 0.06 to 0.13 when the comorbidity variables were added. For total charges, the R^2 increased by 44%, from 0.18 to 0.26, after adding the comorbidity set. For mortality, the comorbidities as a group were not significant. Although the Hosmer–Lemeshow statistic for the logistic regression on all patients did not indicate that the model fit the data, the correspondence between observed and expected values was quite clear over all deciles of risk. The significant departure from model fit was no doubt because of the large sample size, enabling small differences between the observed and expected values to be considered significant.

A number of the comorbidities had substantial effects on the outcomes. Ten comorbidities were consistently influential across all three outcomes, independently increasing resource use by 25% or more or increasing the odds of dying by 50% or more. These were: congestive heart failure, pulmonary circulation disorders, paralysis, other neurological disorders, AIDS, lymphoma, metastatic cancer, coagulopathy, weight loss, and fluid and electrolyte disorders.

Ten Homogeneous Subgroups

As for the heterogeneous population, the relationship between comorbidities and outcomes was strong and was more consistent for resource use than for mortality. Appendixes A through C summarize the effects of comorbidities on length

TABLE 2. Characteristics of Study Population of Adult, Nonmaternal Patients Discharged From 438 California Hospitals in 1992 (n = 1,779,167)

Age (mean) (yr)	57.1
Length of stay (mean) (days)	6.0
Total hospital charges (mean) ($)	14,677
Died in hospital (%)	4.7
Race (%)	
White (referent category)	70.9
Black	8.9
Other	20.2
Male (%)	47.9
Emergency admission (%)	17.0
Insurance status (%)	
Private (referent category)	37.0
Medicare	40.3
Medicaid	12.0
Self-pay	4.0
Other payers	6.7
Surgical DRG[a]	37.5
Complications[b]	6.3

[a]The following are surgical DRGs: 1–8, 36–42, 49–63, 75–77, 103–120, 146–171, 191–201, 209–234, 257–270, 285–293, 302–315, 334–345, 353–365, 370–371, 374–375, 377, 381, 392–394, 400–402, 406–408, 415, 424, 439–443, 458–459, 461, 468, 471–472, 476–486, 488, 491.

[b]The following ICD-9-CM codes define complications: 349.0, 349.1, 429.4, 512.1, 519.0, 564.2, 564.3, 564.4, 569.6, 579.3, 909.3, 995.4, 997.0, 997.1, 997.2, 997.3, 997.4, 997.5, 997.60, 997.61, 997.62, 997.69, 997.9, 998.0, 998.1, 998.2, 998.3, 998.4, 998.5, 998.6, 998.7, 998.8, 998.81, 998.82, 998.89, 998.9, 999.0, 999.1, 999.2, 999.3, 999.4, 999.5, 999.6, 999.7, 999.8, 999.9.

of stay (Appendix A), hospital charges (Appendix B), and in-hospital mortality (Appendix C) for the 10 diagnosis subgroups based on analyses parallel to the multivariate analyses of the entire population. A P value of 0.01 was used for all analyses on the homogeneous subgroups. For example, Appendix A illustrates the number of diagnosis subgroups for which the comorbidities were positively or negatively related to the length of stay at the 0.05 level of statistical significance. In Appendix A, congestive heart failure (CHF) was associated with a statistically significant increase in the length of stay for all 10 diagnosis subgroups; in contrast, lymphoma was associated with a statistically significant increase in length of stay for only two diagnosis subgroups: biliary tract disease and pneumonia. (Regression coefficients for comorbidity effects within the 10 subgroups are available from the authors).

The appendixes show consistently positive and statistically significant results in nearly all of the length-of-stay and charge analyses, but in the mortality analysis, more comorbidities were negatively related to in-hospital death. Again, the probable coding bias of lower mortality being associated with certain comorbidities occurred for many of the same comorbidities as in the analysis of the heterogeneous population. Comorbidities that significantly increased length of stay and charges in many subgroups but were associated with lower mortality in at least one subgroup were: hypertension, obesity, deficiency anemias, and depression. Compared with the effect of comorbidities on mortality in the heterogeneous population (where only 72% of the statistically significant effects were positive), considerably more were positive for the disease-specific tests (85% of the comorbidity tests).

Twenty-six of the 30 comorbidities had at least one diagnosis subgroup for which the specific comorbidity had an effect of 25% or more on resource use or had an effect of increasing the probability of death by 50% or more compared with those without the comorbidity. A few comorbidities were consistently strong predictors of outcomes for the diagnosis subgroups: congestive heart failure, diabetes with complications, renal failure, coagulopathy, weight loss, fluid and electrolyte disorders, blood loss anemia (except for mortality prediction), and deficiency anemia (except for mortality prediction). Of all comorbidities examined, seven were statistically significant predictors in six or more subgroups for all three outcome measures. Those seven comorbidities were: congestive heart failure, cardiac arrhythmias, other neurologic disorders, renal failure, coagulopathy, weight loss, and fluid and electrolyte disorders.

As in the heterogeneous analyses, the effects of the comorbidities as a group on the resource use measures were evaluated for each disease group. The R^2 values for the length-of-stay regressions increased, on average, by 37% when all comorbidities were included, with the increase as high as 75% for pneumonia patients (R^2 from 0.08 to 0.14) and as low as 12% (R^2 from 0.34 to 0.38) for diverticulosis patients. The increase in R^2 for the total-charge regressions were, on average, 49%

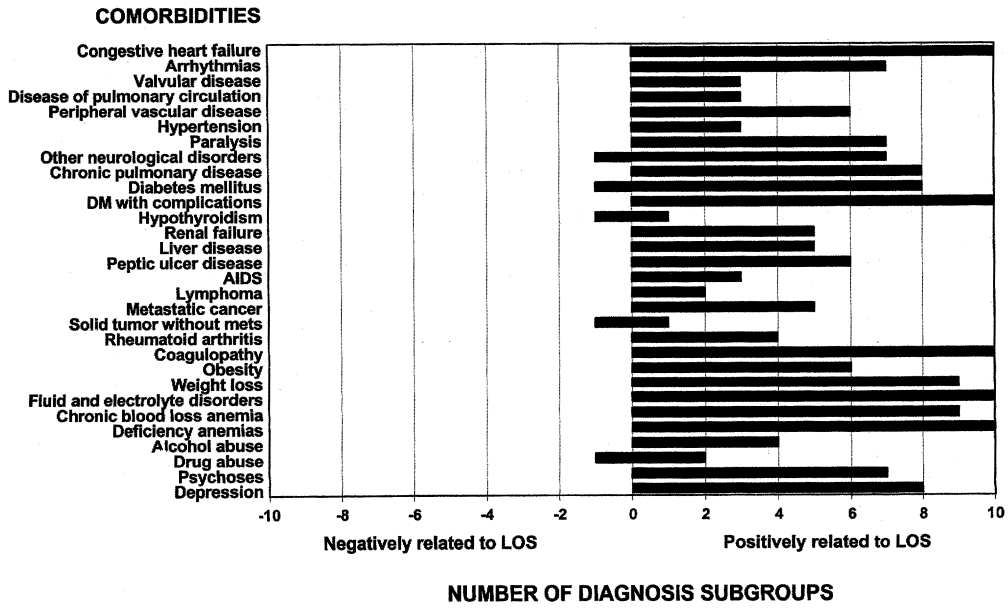

FIG. 1. Number of diagnosis subgroups for which the comorbidities are positively or negatively related to: Length of Stay.

higher when all comorbidities were included, ranging from a 35% increase for low back pain (R^2 from 0.29 to 0.39) to a 200% increase for breast cancer (R^2 from 0.04 to 0.12). The Hosmer–Lemeshow statistic for each diagnosis subgroup indicated a good model fit.

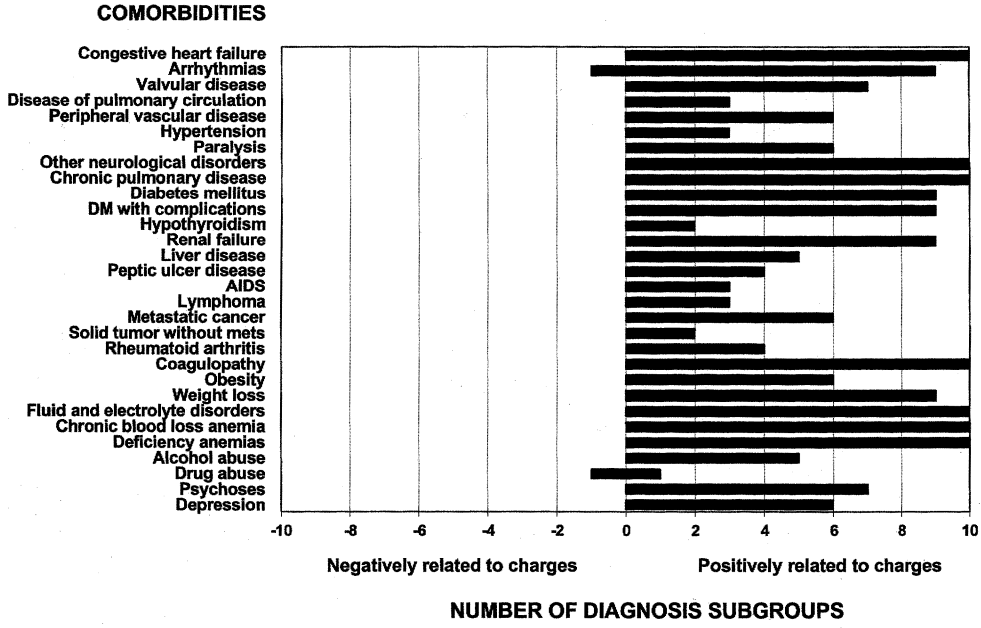

FIG. 2. Number of diagnosis subgroups for which the comorbidities are positively or negatively related to: Total Hospital Charges.

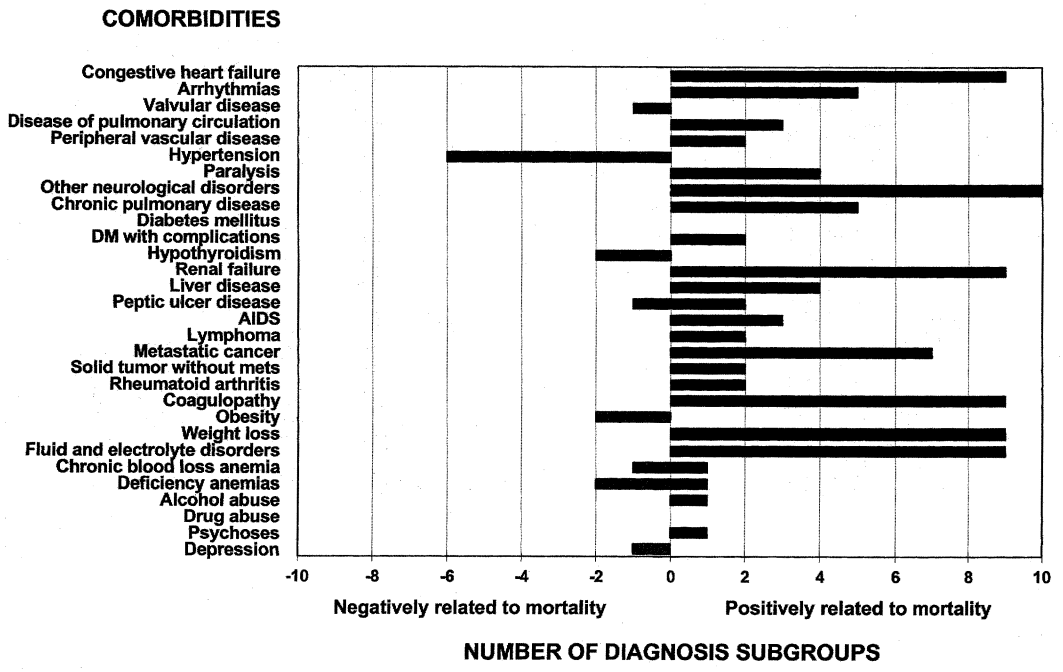

FIG. 3. Number of diagnosis subgroups for which the comorbidities are positively or negatively related to: In-Hospital Mortality.

Discussion

Controlling for preexisting clinical conditions, or comorbidities, is of interest in all types of health care studies. This study identified a set of diagnoses that may represent comorbidities in administrative inpatient datasets. The impacts of the comorbidities on the outcomes of interest were strong. The effects were generally consistent across the outcomes but with varying impacts by specific comorbidities and for specific homogeneous patient groups.

Limitations

The present study has some limitations, which include: (1) the reliance on administrative data, (2) less precise results in predicting mortality, and (3) incomplete validation.

First, this work relies on administrative data, which are never complete or detailed enough to provide a clinically precise method for identifying comorbidities. The most important shortcoming of administrative data, for this study, is that it is not possible to identify when a condition became apparent. Complete claims databases that include inpatient and ambulatory care claims with time can identify the timing of diagnoses and offer a way to distinguish between complications and comorbidities. With such information, the distinction between comorbidities and complications is much more clear, although still subject to the vagaries of coding and data accuracy. Our method attempted to distinguish between comorbidities and complications by making a number of assumptions. Although these assumptions may hold true for many cases, it was still an imperfect procedure, the limitations of which must be taken into account when using this set of comorbidities in other studies.

Another limitation of administrative data is that the distinction between the principal diagnosis and secondary diagnoses may be arbitrary and may be based on nonclinical decisions. Although there are clear coding guidelines for what constitutes a principal diagnosis—the diagnosis that after evaluation is determined to be the main reason for the hospitalization—these guidelines may

TABLE 3. Relationship Between Number of Comorbidities and Outcomes

	Number of Comorbidities			
Characteristic	0	1	2	3+
% of observations	40.3	24.9	17.0	17.8
Age (mean, yr)[a]	48.5	58.8	63.9	68.1
Length of stay (mean, days)[a]	4.4	5.9	7.1	8.9
Total charges (mean, $)[a]	10,944	14,264	16,596	21,882
Died in the hospital (%)[a]	1.6	3.7	6.2	11.7

[a]All differences are significant at $P = 0.05$ based on bivariate analysis of variance.

be subordinated to other incentives such as the desire to maximize reimbursement.

A third limitation of administrative data that affected this study pertains to the ability to distinguish between independent comorbidities and conditions that are directly related to the principal diagnosis. It is possible that our estimates overstated the contribution of comorbidities to explaining resource use and outcomes because of the difficulty of distinguishing between these two concepts. Although we attempted to address this issue by eliminating conditions that were related directly to the DRG for the case, this method was still an approximation. This approach provided a relatively conservative method of defining comorbidities, that is, we tried to avoid identifying a condition as a comorbidity when there was a reasonable likelihood that it represented a further specification of the principal diagnosis or a complication rather than a separate condition.

Despite these limitations, statistical results (ours and others) have supported the use of administrative data for many types of studies, and the present work attempts to aid future studies using these data by offering a reasonable, more comprehensive method for controlling for potentially preexisting comorbidities. Extracting clinical detail for more precise measurement is not only prohibitively expensive on the scale needed to develop a general comorbidity tool, but also would not be broadly applicable to the administrative data that are in common use today.

Second, the final comorbidity list predicted mortality less consistently than it predicted the resources used to treat patients. One possible explanation for this is the awareness of hospital administrators of the importance of record keeping for reimbursement of hospital services. Specifically, the hospital billing clerk is trained (or perhaps uses commercial software) to abstract and record diagnoses that augment reimbursements. The same incentives do not exist for diagnoses that predict mortality, and these may or may not be coded, especially when there are numerous complications already listed on a summary record for a patient who died in the hospital.

Another explanation of the weaker mortality results is that the potential comorbidities that we ruled out because they also were potential complications of treatment—pneumonia, pleural effusion, urinary tract infection, cardiac arrest, cardiogenic shock, and respiratory failure—are indeed important predictors of mortality. Other work has suggested strong associations between several of these omitted diagnoses and mortality of the patient.[29] Thus, without clinical information in addition to the typical administrative record, this method may underestimate the impact of certain comorbidities on mortality.

Further, death in the hospital is a rare event; only 4.7% of the study records were of deceased patients. Combinations of principal diagnoses and comorbidities as well as specific test results and vital signs may be necessary to predict the relatively rare event of hospital death. This may explain why the comorbidities were better predictors of mortality in the diagnosis-specific analyses than in the heterogeneous group.

The third issue is the development and testing of the comorbidity set on the same data. This approach means that the improvement in the percentage of explained variance that was shown in the final regressions with the comorbidities may be overstated. Users may not obtain the same improvement in R^2 (with exclusion and then inclusion of the comorbidity set) that was achieved af-

TABLE 4. Effects of Specific Comorbidities on Outcomes Controlling for Demographic, Insurance, and Other Clinical Factors of Adult, Nonmaternal Patients Who Were Hospitalized in California in 1992 (n = 1,779,167)

Comorbidity	% of Cases With Comorbidity	Length of Stay (transformed coefficients)[a]	Hospital Charges (transformed coefficients)[a]	In-hospital Mortality (odds ratio)[b]
1. Congestive heart failure	4.0	1.35	1.35	2.3
2. Cardiac arrhythmias	6.8	1.12	1.13	1.4
3. Valvular disease	1.8	1.06	1.06	0.7
4. Pulmonary circulation disorders	0.3	1.28	1.48	1.9
5. Peripheral vascular disorders	2.6	1.11	1.16	1.2
6. Hypertension	17.9	NS	1.06	0.6
7. Paralysis	1.8	1.82	1.60	1.7
8. Other neurological disorders	2.7	1.26	1.27	2.8
9. Chronic pulmonary disease	9.9	1.19	1.25	1.2
10. Diabetes, uncomplicated	7.8	1.06	1.13	NS
11. Diabetes, complicated	4.1	1.19	1.19	1.1
12. Hypothyroidism	2.7	1.07	1.06	0.7
13. Renal failure	3.3	0.96	1.12	2.1
14. Liver disease	1.3	1.19	1.17	1.9
15. Peptic ulcer disease excluding bleeding	0.8	1.15	1.13	0.8
16. Acquired immune deficiency syndrome (AIDS)	0.4	1.30	1.45	3.2
17. Lymphoma	0.5	1.25	1.32	1.8
18. Metastatic cancer	2.4	1.43	1.32	3.1
19. Solid tumor without metastasis	6.0	NS	NS	NS
20. Rheumatoid arthritis/collagen vascular diseases	1.2	1.17	1.16	NS
21. Coagulopathy	1.5	1.39	1.75	4.1
22. Obesity	2.3	1.09	1.13	0.5
23. Weight loss	1.1	1.73	1.70	3.2
24. Fluid and electrolyte disorders	13.3	1.27	1.38	2.7
25. Blood loss anemia	1.6	1.23	1.22	0.9
26. Deficiency anemias	7.3	1.30	1.30	NS
27. Alcohol abuse	2.9	1.09	1.09	1.1
28. Drug abuse	1.5	1.35	1.13	NS
29. Psychoses	1.4	1.25	1.14	1.2
30. Depression	1.5	1.27	1.17	0.6
R^2 including all comorbidity variables	—	0.13	0.26	—
R^2 excluding all comorbidity variables	—	0.06	0.18	—

NS, not significant.

[a]Significant at P = 0.0001; antilogarithms of coefficients from ordinary least squares regressions of the logarithm of the dependent variable run on variables specified in the model.

[b]Significant at P = 0.05.

ter developing and then testing the final list on the same data set. The reason that we did not split the data set between development and testing is that we wanted to avoid the weaknesses of earlier work that relied on too few cases and that addressed comorbidities for only a few homogeneous groups of patients. Nonetheless, the improvements in R^2 were substantial, and reductions in the explanatory power of these comorbidity measures will be small in comparison. We also believe that real proof of the value of a measurement tool comes from validation by independent researchers.

Advantages of this Comorbidity Set

This comorbidity set attempts to address weaknesses of previous work on comorbidities and extends this work to take advantage of the information available in administrative data. The identification of these comorbidities involved a comprehensive approach including a survey of the literature and the ICD-9-CM manual, which allowed us to begin with a very broad, inclusive list of comorbidities, many of which have not been explored previously, and to consider the complexity of the ICD-9-CM coding system in defining the comorbidity set. Many of these conditions could be complications of medical care; many of them could be indicators of the severity of the principal diagnosis. The methods described here attempted to control for this, but the realities of administrative data mean that it was not possible to easily compartmentalize all conditions into the separate concepts defined earlier: the primary reason for admission, the severity of this principal diagnosis, complications that result from medical care, and comorbidities. There is little evidence on which to base judgments of whether a condition is more often a comorbidity or more often a complication or more often a further definition of the principal diagnosis. In the absence of such information, we provide a method that allows for a flexible approach to measuring comorbidities, a method that should be tempered by the clinical judgment of researchers. For example, users may elect to exclude some comorbidities depending on the specific patient population that is being studied.

The measures were developed on a large, state population from multiple institutions that was representative of adult hospitalized patients in general and that allowed us to assess a broad range of comorbidities. The measures were constructed to be used on a heterogeneous patient population with disparate principal diagnoses as well as on specific, homogeneous diagnosis groups. Most other studies have tested comorbidities on limited sets of patients.[3,4,7–12,15,17,19,29,31,34] The resulting comorbidity measures are applicable to large administrative datasets, which are now widely available at the national, state, health plan, hospital, and, in some cases, physician levels. The comorbidity measures were developed to predict outcomes most commonly evaluated in research on hospital services—length of stay, charges, and in-hospital death.

Comorbidity measures that have relied exclusively on administrative information generally have had a problem distinguishing comorbidities from complications and from the underlying reason for the hospitalization.[6] For example, the Charlson Comorbidity Index minimized this problem by concentrating only on chronic comorbid conditions, at the expense of acute illnesses.[1,19] Our method provides a DRG screen, which allowed us to include not only chronic but also potentially acute illnesses, such as congestive heart failure and cardiac arrhythmias, as comorbidities without including them when they were related to the reason for the hospitalization.

The Charlson list of comorbidities has been used by many researchers as a method to control for the effects of comorbidities in various analyses. Our list adds a number of comorbidities to the Charlson list and drops several that were never related to outcomes or that were conceptually inappropriate. Three of the newly identified comorbidities had strong, consistent results across the outcomes measured: coagulopathy, weight loss, and fluid and electrolyte disorders. The remaining new comorbidities were important predictors for some outcomes or specific diseases.

Other comorbidity lists must be revised by users to explicitly omit comorbidities related to the principal diagnosis under study. For example, Deyo et al[19] had to omit musculoskeletal comorbidities from Charlson's set of comorbidities when they were studying patients undergoing lumbar spine surgery. With our method, the user can apply the comorbidity list with its DRG screen without further refinement because the DRG screen eliminates comorbidities related to the principal diagnosis. Thus, our measure is generally applicable to studies of all types of diseases.

Rather than attempting to derive a summary measure of comorbidities, we chose to retain comorbidities as separate, independent measures because individual comorbidities will be irrelevant for some diseases and are likely to influence the outcomes of different diseases and treatments differently. This was apparent in our work and in studies by others.[6,18,30] Our tool is a flexible measure that will allow investigators to examine the varying impact of different comorbidities on different outcome measures, a limitation of approaches that rely on an index.

Despite the limitations identified earlier, this work represents an improvement in the methods available today for measuring comorbidities in large administrative datasets. We hope this measure will be tested and validated through applications by other researchers to other data systems, to other conditions and diseases, and to other populations, such as the elderly.

Acknowledgments

The authors thank Sandy Smoot of Social and Scientific Systems for her work in statistical programming and Nadine Stuart and Emuella Flood for their help in preparing the manuscript.

References

1. **Charlson ME, Pompei P, Ales KL, MacKenzie CR.** A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Chronic Disease 1987;40:373.

2. **Iezzoni LI, Ash AS, Shwartz M, Daley J, Hughes JS, Mackiernan.** Predicting who dies depends on how severity is measured: Implications for evaluating patient outcomes. Ann Intern Med 1995;123:763.

3. **Matsui K, Goldman L, Johnson PA, Kuntz KM, Cook EF, Lee TH.** Comorbidity as a correlate of length of stay for hospitalized patients with acute chest pain. J Gen Intern Med 1996;11:262.

4. **Iezzoni LI, Shwartz M, Ash AS, Hughes JS, Daley J, Mackiernan YD.** Severity measurement methods and judging death rates for pneumonia. Med Care 1996;34:11.

5. **Fowles JB, Weiner JP, Knutson D, Fowler E, Tucker AM, Ireland M.** Taking health status into account when setting capitation rates. JAMA 1996;276:1316.

6. **Hughes JS, Iezzoni LI, Daley JD, Greenberg LG.** How severity measures rate hospitalized patients. J Gen Intern Med 1996;11:303.

7. **Landon B, Iezzoni LI, Ash AS, Daley J, Hughes JS, Mackiernan YD.** Judging hospitals by severity-adjusted mortality rates: The case of CABG surgery. Inquiry 1996;33:155.

8. **Litwin MS, Kahn KL, Reccius N.** Why do sicker patients cost more? A charge-based analysis of patients undergoing prostatectomy. Journal of Urology 1993;149:84.

9. **Steiner CA, Bass EB, Talamini MA, Pitt HA, Steinberg EP.** Surgical rates and operative mortality for open and laparoscopic cholecystectomy in Maryland. N Engl J Med 1994;330:403.

10. **Greenfield S, Apolone G, McNeil BJ, Cleary PD.** The importance of coexistent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Med Care 1993;31:141.

11. **Bennett CL, Greenfield S, Aronow H, Ganz P, Vogelzang NJ, Elashoff RM.** Patterns of care related to age of men with prostate cancer. Cancer 1991;67:2633.

12. **Kaplan MH, Feinstein AR.** The importance of classifying initial comorbidity in evaluating the outcome of diabetes mellitus. Journal of Chronic Disease 1974;27:387.

13. **Keeler EB, Kahn KL, Draper D, et al.** Changes in sickness at admission following the introduction of the prospective payment system. JAMA 1990;264:1962.

14. **Naessens JM, Leibson CL, Krishan I, Ballard DJ.** Contribution of a measure of disease complexity (COMPLEX) to prediction of outcome and charges among hospitalized patients. Mayo Clin Proc 1992;67:1140.

15. **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.

16. **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.

17. **D'Hoore W, Sicotte C, Tilquin C.** Risk adjustment in outcome assessment: The Charlson Comorbidity Index. Methods Inf Med 1993;32:382.

18. **Hartz AJ, Kuhn EM, Kayser KL, Pryor DP, Green R, Rimm AA.** Assessing providers of coronary revascularization: A method for peer review organizations. Am J Public Health 1992;82:1631.

19. **Roos LL, Sharp SM, Cohen MM, Wajda A.** Risk adjustment in claims-based research: The search for efficient approaches. J Clin Epidemiol 1989;42:1193.

20. **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.

21. **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.

22. **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.

23. **Health Care Financing Administration.** International Classification of Diseases, 9th Revision, Clinical Modification, Volumes 1–3, DHHS Publication No. (PHS) 80–1260.

24. **3M.** DRGs: Diagnosis Related Groups, Definitions Manual, Version 10.0 (effective October 1, 1992).

25. **Gonnella JS, Hornbrook MC, Louis DZ.** Staging of disease: A casemix measurement. JAMA 1984; 251:637.

26. **Jencks SF, Williams DK, Kay TL.** Assessing hospital-associated deaths from discharge data: The role of length of stay and comorbidities. JAMA 1988;260:2240.

27. **Iezzoni LI, Foley SM, Daley J, Hughes J, Fisher ES, Heeren T.** Comorbidities, complications, and coding bias: Does the number of diagnosis codes matter in predicting in-hospital mortality? JAMA 1992;267: 2197.

28. **Hannan EL, Kilburn H, Lindsey ML, Lewis R.** Clinical versus administrative data bases for CABG surgery: Does it matter? Med Care 1992;30:892.

29. **Iezzoni LI, Heeren T, Foley SM, Daley J, Hughes J, Coffman GA.** Chronic conditions and risk of in-hospital death. Health Serv Res 1994;29:435.

30. **Krakauer H, Bailey RC, Skellan KJ, et al.** Evaluation of the HCFA model for analysis of mortality following hospitalization. Health Serv Res 1992; 27:317.

31. **McGovern PG, Burke GL, Sprafka JM, Xue S, Folsom AR, Blackburn H.** Trends in mortality, morbidity, and risk factor levels for stroke from 1960 through 1990. JAMA 1992;268:753.

32. **DesHarnais S.** Current uses of large data sets to assess the quality of providers. Construction of risk-adjusted indexes of hospital performance. Int J Technol Assess Health Care 1990;6:229.

Appendix A.
Effects of Comorbidities on Length of Stay for 10 Disease Subgroups: Transformed Coefficients (*P* Values)

Comorbidity	Breast Cancer	AMI	Asthma	Appendicitis	Hernia	Diverticulosis	Biliary Tract Disease	Low Back Pain	Pneumonia	Complicated Diabetes
1. Congestive heart failure	1.25 (0.0001)	1.55 (0.0001)	1.23 (0.0001)	1.39 (0.0001)	1.30 (0.0001)	1.20 (0.0001)	1.48 (0.0001)	1.20 (0.0001)	1.20 (0.0001)	1.27 (0.0001)
2. Cardiac arrhythmias	1.08 (0.0034)	NS	1.16 (0.0001)	NS	NS	1.09 (0.0001)	1.13 (0.0001)	1.09 (0.0001)	1.07 (0.0001)	1.06 (0.0013)
3. Valvular disease	—	1.26 (0.0112)	NS	NS	NS	NS	1.12 (0.0001)	NS	1.04 (0.0035)	NS
4. Pulmonary circulation disorders	—	NS	1.39 (0.0001)	NS	NS	NS	1.30 (0.0067)	NS	1.17 (0.0001)	NS
5. Peripheral vascular disorders	NS	1.06 (0.0006)	NS	1.21 (0.0365)	1.22 (0.0002)	1.15 (0.0001)	1.11 (0.0027)	1.09 (0.0054)	NS	1.21 (0.0001)
6. Hypertension	NS	1.03 (0.0001)	1.05 (0.0003)	NS	NS	NS	NS	1.05 (0.0001)	NS	NS
7. Paralysis	1.35 (0.0001)	1.11 (0.0008)	1.19 (0.0106)	NS	—	NS	1.39 (0.0001)	1.39 (0.0001)	1.14 (0.0001)	1.14 (0.0006)
8. Other neurological disorders	1.21 (0.0014)	0.90 (0.0001)	1.23 (0.0001)	NS	NS	1.15 (0.0005)	1.22 (0.0001)	1.20 (0.0001)	1.14 (0.0001)	1.22 (0.0001)
9. Chronic pulmonary disease	1.06 (0.0166)	1.15 (0.0001)	1.48 (0.0001)	NS	1.15 (0.0001)	1.12 (0.0001)	1.17 (0.0001)	1.09 (0.0001)	1.09 (0.0001)	1.08 (0.0001)
10. Diabetes, uncomplicated	NS	1.05 (0.0001)	1.09 (0.0001)	1.15 (0.0001)	1.11 (0.0005)	1.07 (0.0003)	1.09 (0.0001)	1.08 (0.0001)	1.03 (0.0037)	0.87 (0.0041)
11. Diabetes, complicated	1.21 (0.0002)	1.16 (0.0001)	1.20 (0.0001)	1.20 (0.0008)	1.27 (0.0001)	1.12 (0.0004)	1.27 (0.0001)	1.17 (0.0001)	1.09 (0.0001)	1.16 (0.0001)
12. Hypothyroidism	NS	1.09 (0.0001)	NS	—	—	1.06 (0.0228)	0.93 (0.0077)	NS	1.03 (0.0268)	NS
13. Renal failure	1.25 (0.0286)	1.09 (0.0001)	NS	1.31 (0.0044)	1.17 (0.0024)	1.20 (0.0001)	1.28 (0.0001)	NS	NS	NS
14. Liver disease	NS	1.21 (0.0015)	NS	1.51 (0.0003)	1.12 (0.0195)	1.12 (0.0357)	1.21 (0.0001)	NS	1.12 (0.0001)	1.12 (0.0063)
15. Peptic ulcer disease excluding bleeding	NS	1.14 (0.0003)	1.22 (0.0001)	1.26 (0.0139)	NS	—	1.27 (0.0001)	—	1.13 (0.0001)	1.16 (0.0110)
16. Acquired immune deficiency syndrome (AIDS)	—	NS	—	—	—	—	1.86 (0.0001)	—	1.21 (0.0001)	1.39 (0.0003)
17. Lymphoma	—	NS	—	—	NS	NS	1.42 (0.0001)	NS	1.13 (0.0001)	NS
18. Metastatic cancer	1.95 (0.0001)	NS	NS	NS	1.39 (0.0001)	1.40 (0.0001)	1.32 (0.0001)	NS	1.07 (0.0001)	1.14 (0.0380)

19. Solid tumor without metastasis	NS	NS	NS	NS	NS	0.93 (0.0011)	1.12 (0.0001)	NS	NS	NS
20. Rheumatoid arthritis/collagen vascular diseases	—	NS	1.14 (0.0102)	—	NS	1.09 (0.0120)	1.15 (0.0006)	1.07 (0.0179)	1.07 (0.0002)	NS
21. Coagulopathy	1.73 (0.0001)	1.20 (0.0001)	1.42 (0.0001)	1.22 (0.0014)	1.55 (0.0001)	1.14 (0.0052)	1.43 (0.0001)	1.28 (0.0001)	1.26 (0.0001)	1.30 (0.0001)
22. Obesity	1.12 (0.0100)	1.04 (0.0427)	1.09 (0.0001)	1.12 (0.0247)	—	1.12 (0.0003)	1.09 (0.0001)	1.09 (0.0001)	NS	1.11 (0.0001)
23. Weight loss	NS	1.46 (0.0001)	1.58 (0.0001)	1.86 (0.0001)	2.20 (0.0001)	1.58 (0.0001)	1.77 (0.0001)	1.42 (0.0015)	1.39 (0.0001)	1.60 (0.0001)
24. Fluid and electrolyte disorders	1.43 (0.0001)	1.15 (0.0001)	1.17 (0.0001)	1.45 (0.0001)	1.68 (0.0001)	1.17 (0.0001)	1.68 (0.0001)	1.28 (0.0001)	1.12 (0.0001)	1.12 (0.0001)
25. Blood loss anemia	1.48 (0.0001)	1.30 (0.0001)	1.49 (0.0001)	1.52 (0.0037)	1.52 (0.0001)	NS	1.62 (0.0001)	1.40 (0.0001)	1.40 (0.0001)	1.45 (0.0001)
26. Deficiency anemias	1.52 (0.0001)	1.22 (0.0001)	1.11 (0.0004)	1.30 (0.0001)	1.25 (0.0001)	1.17 (0.0001)	1.45 (0.0001)	1.35 (0.0001)	1.16 (0.0001)	1.27 (0.0001)
27. Alcohol abuse	1.26 (0.0381)	1.11 (0.0026)	NS	1.23 (0.0065)	1.23 (0.0026)	1.12 (0.0351)	1.34 (0.0001)	NS	NS	NS
28. Drug abuse	—	NS	NS	NS	—	1.40 (0.0060)	1.20 (0.0189)	1.67 (0.0001)	0.95 (0.0138)	NS
29. Psychoses	1.35 (0.0001)	NS	1.19 (0.0001)	—	NS	1.20 (0.0008)	1.46 (0.0001)	1.27 (0.0001)	1.17 (0.0001)	1.25 (0.0001)
30. Depression	1.27 (0.0025)	1.14 (0.0013)	1.25 (0.0001)	1.52 (0.0010)	—	NS	1.23 (0.0001)	1.48 (0.0001)	1.16 (0.0001)	1.21 (0.0001)
R^2 including comorbidity variables	0.20	0.13	0.17	0.26	0.22	0.38	0.27	0.14	0.14	0.20
R^2 excluding comorbidity variables	0.14	0.09	0.13	0.23	0.16	0.34	0.20	0.10	0.08	0.14

AMI, acute myocardial infarction; NS, not significant; — indicates too few cases had this comorbidity in this model to reliably estimate a coefficient.

Appendix B
Effects of Comorbidities on Hospital Charges for 10 Disease Subgroups: Transformed Coefficients (*P* Values)

Comorbidity	Breast Cancer	AMI	Asthma	Appendicitis	Hernia	Diverticulosis	Biliary Tract Disease	Low Back Pain	Pneumonia	Complicated Diabetes
1. Congestive heart failure	1.27 (0.0001)	1.16 (0.0001)	1.27 (0.0001)	1.63 (0.0001)	1.28 (0.0001)	1.27 (0.0001)	1.42 (0.0001)	1.22 (0.0001)	1.28 (0.0001)	1.30 (0.0001)
2. Cardiac arrhythmias	1.11 (0.0001)	0.86 (0.0012)	1.26 (0.0001)	1.11 (0.0012)	1.11 (0.0001)	1.19 (0.0001)	1.15 (0.0001)	1.15 (0.0001)	1.21 (0.0001)	1.21 (0.0001)
3. Valvular disease	—	1.32 (0.0020)	1.11 (0.0052)	1.13 (0.0319)	NS	1.11 (0.0027)	1.15 (0.0001)	1.08 (0.0016)	1.08 (0.0001)	1.11 (0.0079)
4. Pulmonary circulation disorders	—	NS	1.54 (0.0001)	NS	NS	NS	1.38 (0.0001)	1.43 (0.0191)	1.31 (0.0001)	NS
5. Peripheral vascular disorders	NS	1.05 (0.0016)	NS	1.23 (0.0099)	1.16 (0.0020)	1.15 (0.0001)	1.06 (0.0251)	1.14 (0.0001)	NS	1.22 (0.0001)
6. Hypertension	1.03 (0.0492)	1.02 (0.0179)	1.05 (0.0001)	NS	1.05 (0.0024)	NS	1.02 (0.0465)	1.06 (0.0001)	NS	NS
7. Paralysis	1.23 (0.0007)	NS	1.19 (0.0095)	NS	—	NS	1.20 (0.0001)	1.35 (0.0001)	1.17 (0.0001)	1.19 (0.0001)
8. Other neurological disorders	1.13 (0.0130)	1.20 (0.0001)	1.36 (0.0001)	1.23 (0.0004)	1.15 (0.0117)	1.19 (0.0001)	1.20 (0.0001)	1.17 (0.0001)	1.23 (0.0001)	1.40 (0.0001)
9. Chronic pulmonary disease	1.07 (0.0003)	1.15 (0.0001)	2.27 (0.0001)	1.14 (0.0001)	1.21 (0.0001)	1.20 (0.0001)	1.15 (0.0001)	1.15 (0.0001)	1.19 (0.0001)	1.11 (0.0001)
10. Diabetes, uncomplicated	1.06 (0.0042)	1.05 (0.0001)	1.09 (0.0001)	1.17 (0.0001)	1.11 (0.0002)	1.09 (0.0001)	1.07 (0.0001)	1.11 (0.0001)	1.04 (0.0001)	—
11. Diabetes, complicated	1.13 (0.0088)	1.16 (0.0001)	1.19 (0.0001)	1.17 (0.0006)	1.20 (0.0001)	1.09 (0.0084)	1.17 (0.0001)	1.20 (0.0001)	1.13 (0.0001)	—
12. Hypothyroidism	NS	1.06 (0.0057)	NS	—	—	1.09 (0.0032)	NS	1.04 (0.0291)	1.03 (0.0426)	NS
13. Renal failure	NS	1.16 (0.0001)	1.23 (0.0003)	1.46 (0.0001)	1.31 (0.0001)	1.34 (0.0001)	1.39 (0.0001)	1.19 (0.0004)	1.16 (0.0001)	1.14 (0.0001)
14. Liver disease	1.20 (0.0326)	1.15 (0.0197)	NS	1.67 (0.0001)	1.15 (0.0022)	1.17 (0.0075)	1.17 (0.0001)	NS	1.17 (0.0001)	NS
15. Peptic ulcer disease excluding bleeding	NS	NS	1.17 (0.0003)	1.35 (0.0003)	NS	—	1.19 (0.0001)	—	1.09 (0.0022)	1.13 (0.0303)
16. Acquired immune deficiency syndrome (AIDS)	—	NS	—	—	—	—	2.08 (0.0001)	—	1.38 (0.0001)	1.35 (0.0010)
17. Lymphoma	—	1.20 (0.0175)	—	—	NS	1.28 (0.0079)	1.26 (0.0001)	NS	1.26 (0.0001)	NS
18. Metastatic cancer	1.82 (0.0001)	NS	NS	NS	NS	1.46 (0.0001)	1.30 (0.0001)	1.22 (0.0002)	1.15 (0.0001)	1.21 (0.0021)

19. Solid tumor without metastasis	—	NS	NS	NS	1.06 (0.0219)	NS	1.07 (0.0001)	NS	1.03 (0.0138)	NS
20. Rheumatoid arthritis/collagen vascular diseases	NS	NS	NS	—	NS	1.17 (0.0001)	1.15 (0.0001)	1.09 (0.0013)	1.09 (0.0001)	NS
21. Coagulopathy	2.01 (0.0001)	1.38 (0.0001)	1.51 (0.0001)	1.32 (0.0014)	1.28 (0.0020)	1.43 (0.0001)	1.55 (0.0001)	1.40 (0.0001)	1.63 (0.0001)	1.55 (0.0001)
22. Obesity	1.12 (0.0021)	NS	1.11 (0.0001)	1.16 (0.0009)	—	1.14 (0.0001)	1.05 (0.0002)	1.11 (0.0001)	NS	NS
23. Weight loss	NS	1.38 (0.0001)	1.67 (0.0001)	1.84 (0.0001)	2.03 (0.0001)	1.60 (0.0001)	1.68 (0.0001)	1.38 (0.0016)	1.51 (0.0001)	1.54 (0.0001)
24. Fluid and electrolyte disorders	1.45 (0.0001)	1.30 (0.0001)	1.30 (0.0001)	1.51 (0.0001)	1.65 (0.0001)	1.23 (0.0001)	1.45 (0.0001)	1.38 (0.0001)	1.27 (0.0001)	1.25 (0.0001)
25. Blood loss anemia	1.48 (0.0001)	1.25 (0.0001)	1.43 (0.0001)	1.52 (0.0011)	1.75 (0.0001)	1.07 (0.0016)	1.39 (0.0001)	1.60 (0.0001)	1.52 (0.0001)	1.72 (0.0001)
26. Deficiency anemias	1.40 (0.0001)	1.20 (0.0001)	1.16 (0.0001)	1.32 (0.0001)	1.28 (0.0001)	1.17 (0.0001)	1.35 (0.0001)	1.39 (0.0001)	1.22 (0.0001)	1.31 (0.0001)
27. Alcohol abuse	1.23 (0.0278)	1.08 (0.0051)	NS	1.27 (0.0005)	NS	1.17 (0.0028)	1.22 (0.0001)	1.28 (0.0001)	NS	1.07 (0.0233)
28. Drug abuse	—	NS	1.08 (0.0356)	NS	—	NS	1.14 (0.0238)	—	0.94 (0.0137)	1.15 (0.0001)
29. Psychoses	1.23 (0.0010)	NS	1.17 (0.0001)	—	NS	1.22 (0.0005)	1.30 (0.0001)	1.26 (0.0001)	1.16 (0.0001)	1.19 (0.0001)
30. Depression	1.19 (0.0103)	NS	1.19 (0.0001)	1.36 (0.0059)	—	NS	1.17 (0.0001)	—	1.14 (0.0001)	1.13 (0.0003)
R^2 including comorbidity variables	0.12	0.43	0.19	0.27	0.25	0.51	0.31	0.32	0.23	0.36
R^2 excluding comorbidity variables	0.04	0.40	0.13	0.22	0.17	0.46	0.22	0.29	0.13	0.29

AMI, acute myocardial infarction; NS, not significant; — indicates too few cases had this comorbidity in this model to reliably estimate a coefficient.

Appendix C
Effects of Comorbidities on In-hospital Mortality for 10 Disease Subgroups: Odds Ratios (*P* Values)

Comorbidity	Breast Cancer	AMI	Asthma	Appendicitis	Hernia	Diverticulosis	Biliary Tract Disease	Low Back Pain	Pneumonia	Complicated Diabetes
1. Congestive heart failure	6.5 (0.0001)	1.9 (0.0001)	2.2 (0.0001)	3.4 (0.0044)	2.2 (0.0039)	4.9 (0.0001)	3.0 (0.0001)	2.3 (0.0458)	1.5 (0.0001)	2.3 (0.0001)
2. Cardiac arrhythmias	NS	1.8 (0.0010)	1.6 (0.0071)	NS	1.7 (0.0293)	NS	1.4 (0.0192)	2.3 (0.0044)	1.3 (0.0001)	1.5 (0.0002)
3. Valvular disease	—	NS	NS	NS	NS	NS	NS	NS	0.6 (0.0001)	NS
4. Pulmonary circulation disorders	—	NS	3.6 (0.0097)	NS	NS	NS	4.1 (0.0063)	NS	1.5 (0.0001)	NS
5. Peripheral vascular disorders	NS	1.2 (0.0326)	NS	NS	2.5 (0.0245)	NS	NS	NS	1.3 (0.0002)	1.4 (0.0143)
6. Hypertension	NS	0.6 (0.0001)	NS	NS	0.5 (0.0142)	0.5	0.5 (0.0009)	NS	0.5 (0.0001)	0.5 (0.0001)
7. Paralysis	NS	2.1 (0.0001)	NS	6.3 (0.0218)	—	NS	NS	3.2 (0.0065)	1.9 (0.0001)	2.7 (0.0001)
8. Other neurological disorders	4.5 (0.0006)	7.2 (0.0001)	8.3 (0.0001)	4.6 (0.0114)	4.3 (0.0006)	2.7 (0.0074)	2.6 (0.0006)	5.5 (0.0001)	2.3 (0.0001)	4.2 (0.0001)
9. Chronic pulmonary disease	NS	NS	9.2 (0.0001)	NS	1.6 (0.0474)	1.6 (0.0070)	1.5 (0.0067)	2.7 (0.0005)	0.7 (0.0001)	NS
10. Diabetes, uncomplicated	NS	NS	NS	NS	NS	NS	NS	NS	0.9 (0.0592)	NS
11. Diabetes, complicated	NS	1.6 (0.0001)	NS	NS	NS	NS	1.9 (0.0027)	NS	1.1 (0.0318)	NS
12. Hypothyroidism	NS	0.5 (0.0001)	NS	—	—	NS	NS	NS	0.8 (0.0041)	NS
13. Renal failure	3.9 (0.0406)	2.1 (0.0001)	3.0 (0.0037)	6.7 (0.0016)	5.0 (0.0001)	4.0 (0.0001)	3.1 (0.0001)	7.7 (0.0001)	2.2 (0.0001)	2.3 (0.0001)
14. Liver disease	NS	2.5 (0.0001)	NS	NS	2.6 (0.0392)	NS	3.8 (0.0001)	NS	2.0 (0.0001)	2.0 (0.0134)
15. Peptic ulcer disease excluding bleeding	NS	0.4 (0.0001)	NS	11.0 (0.0053)	4.5 (0.0078)	—	NS	—	NS	NS
16. Acquired immune deficiency syndrome (AIDS)	—	NS	—	—	—	—	15.6 (0.0001)	—	2.3 (0.0001)	7.4 (0.0001)
17. Lymphoma	—	NS	—	—	NS	NS	3.5 (0.0084)	NS	1.6 (0.0001)	NS
18. Metastatic cancer	14.7 (0.0001)	1.5 (0.0115)	NS	NS	NS	4.1 (0.0033)	4.2 (0.0001)	10.8 (0.0002)	2.9 (0.0001)	4.0 (0.0001)

19. Solid tumor without metastasis	—	NS	NS	NS	NS	NS	1.7 (0.0057)	NS	1.3 (0.0001)	NS
20. Rheumatoid arthritis/collagen vascular diseases	NS	NS	NS	—	NS	4.2 (0.0001)	2.6 (0.0090)	NS	NS	NS
21. Coagulopathy	2.4 (0.0353)	2.6 (0.0001)	8.9 (0.0001)	10.8 (0.0025)	6.8 (0.0002)	5.6 (0.0001)	7.8 (0.0001)	10.3 (0.0001)	3.3 (0.0001)	2.6 (0.0001)
22. Obesity	NS	0.5 (0.0001)	NS	NS	—	NS	NS	NS	0.5 (0.0001)	0.5 (0.0415)
23. Weight loss	NS	2.7 (0.0001)	6.5 (0.0001)	6.2 (0.0036)	4.1 (0.0016)	3.1 (0.0002)	3.0 (0.0001)	10.9 (0.0035)	2.9 (0.0001)	4.0 (0.0001)
24. Fluid electrolyte disorders	2.4 (0.0001)	2.0 (0.0001)	2.8 (0.0001)	2.3 (0.0198)	4.1 (0.0001)	2.6 (0.0001)	3.1 (0.0001)	3.1 (0.0001)	2.3 (0.0001)	2.1 (0.0001)
25. Blood loss anemia	NS	0.5 (0.0001)	NS	NS	NS	NS	NS	NS	1.3 (0.0086)	1.8 (0.0191)
26. Deficiency anemias	1.8 (0.0267)	0.7 (0.0001)	NS	NS	2.4 (0.0037)	NS	NS	NS	0.8 (0.0001)	NS
27. Alcohol abuse	NS	NS	NS	NS	5.6 (0.0007)	3.0 (0.0244)	NS	NS	NS	NS
28. Drug abuse	—	NS	NS	NS	—	NS	NS	NS	NS	NS
29. Psychoses	NS	NS	NS	—	NS	NS	NS	10.5 (0.0001)	1.2 (0.0311)	NS
30. Depression	NS	0.6 (0.0042)	NS	NS	—	NS	NS	NS	0.8 (0.0247)	NS
Significant Hosmer-Lemeshow statistic										

AMI, acute myocardial infarction; NS, not significant; — indicates too few cases had this comorbidity in this model to reliably estimate a coefficient.