

MEDICAL CARE
Volume 40, Number 8, Supplement, pp IV-26-IV-35
©2002 Lippincott Williams & Wilkins, Inc.

Assessing Comorbidity Using Claims Data An Overview

CARRIE N. KLABUNDE, PhD,* JOAN L. WARREN, PhD,* AND JULIE M. LEGLER, ScD†

Comorbidity, additional disease beyond the condition under study that increases a patient's total burden of illness, is one dimension of health status. For investigators working with observational data obtained from administrative databases, comorbidity assessment may be a useful and important means of accounting for differences in patients' underlying health status. There are multiple ways of measuring comorbidity. This paper provides an overview of current approaches to and issues in assessing comorbidity using claims data, with a particular focus on established indices and the SEER-Medicare database. In addition, efforts to improve measurement of comorbidity using claims data are described, including augmentation of claims data with medical record, patient self-report, or health services utilization data; incorporation of

claims data from sources other than inpatient claims; and exploration of alternative conditions, indices, or ways of grouping conditions. Finally, caveats about claims data and areas for future research in claims-based comorbidity assessment are discussed. Although the use of claims databases such as SEER-Medicare for health services and outcomes research has become increasingly common, investigators must be cognizant of the limitations of comorbidity measures derived from these data sources in capturing and controlling for differences in patient health status. The assessment of comorbidity using claims data is a complex and evolving area of investigation.

Key words: Comorbidity; claims data; administrative data; SEER; Medicare. (Med Care 2002;40[supplement]:IV-26-IV-35)

Administrative databases have become an accepted and commonly used resource in health services and outcomes research studies. In studies using administrative databases, however, the lack of randomization inherent in observational data presents an important need to account for differences in patients' underlying health status. Comorbidity assessment is one means of adjusting for these differences, although it is important to recognize that comorbidity is only one dimension of health status; others include age, gender, func-

tional status, and psychological, cognitive, and psychosocial functioning.¹

Comorbidities are additional diseases beyond the condition under study.^{2,3} They can be chronic diseases or acute illnesses, and increase a patient's total burden of illness.^{1,4} In the general population, the presence of comorbid conditions is related to increasing age, lower educational attainment, and public health insurance,³ and is predictive of longer hospital stays, higher hospital costs, in-hospital mortality, and readmission.⁴⁻⁶

*From the Health Services and Economics Branch, Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland.

†From the Statistical Research and Applications Branch, Surveillance Research Program, Division of

Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland.

Address correspondence and reprint requests to: Carrie N. Klabunde, PhD, Applied Research Program, National Cancer Institute, Executive Plaza North Room 4005, 6130 Executive Boulevard, Bethesda, Maryland, 20892-7344. E-Mail: ck97b@nih.gov

Materials and Methods

Selected Approaches to Assessing Comorbidity Using Claims Data

One population-based study found that 69% of cancer patients aged 40 to 84 years had one or more chronic illnesses in addition to cancer, with African Americans more likely than whites to have comorbidity.⁷ In cancer patients, a greater burden of comorbid illness is predictive of less aggressive cancer treatment⁸⁻¹² as well as early mortality.¹²⁻¹⁴ For studies that use administrative or claims databases, the primary source of comorbidity data comprises the diagnosis code fields in which the reason for the hospitalization or visit, as well as any accompanying conditions that required attention, are recorded. For Medicare Part A inpatient hospitalization files, which contain a summarized record of each hospital inpatient stay, there are up to 10 diagnoses and 10 procedures recorded using ICD-9-CM codes.¹⁵ For Medicare Part B physician claims files, a Common Procedure Terminology (CPT) code¹⁶ is required for each billed procedure, and each CPT code must be accompanied by an ICD-9-CM diagnosis code that explains the reason for the visit. Because the ICD-9-CM coding system contains over 14,000 different diagnosis and procedure codes,¹ researchers are presented with the challenge of determining how best to select comorbid conditions for a given study.¹⁷ One approach is to solicit expert clinical opinion, whereas another is to use statistical techniques to identify "important" comorbidities from the dataset of interest. A major drawback of both approaches, however, is the potential lack of comparability across studies, because expert opinion or statistical methodology may vary from study to study. A third approach involves applying an established comorbidity index or algorithm. Through a review of the literature, this paper provides an overview of current approaches to and issues in assessing comorbidity using claims data, with a particular focus on established indices and the SEER-Medicare database. We do not delineate a prescribed methodology for comorbidity assessment using SEER-Medicare or other claims databases, because this is a complex and evolving area of investigation, and possible methodologies will vary depending on the data that are available and the research question to be addressed. Rather, we describe methodologies developed by others that may have utility with SEER-Medicare data, caveats about using claims data to measure comorbidity, efforts to improve this measurement, and areas for future research in claims-based comorbidity measurement.

The Charlson index¹⁸ is perhaps the most well-known and widely used comorbidity measure.^{17,23} Initially developed in 1987, the Charlson index is a summary measure of 19 comorbid conditions, each of which is assigned a weight according to its potential for influencing mortality. Charlson and colleagues demonstrated the index to predict 1-year all-cause mortality in a cohort of 559 hospitalized medical service patients, 4- and 10-year nonbreast cancer mortality in a cohort of 685 breast cancer patients. The Charlson index was adapted in the early 1990s for use with large administrative databases by two different groups: Deyo et al²⁴ and Dartmouth-Manitoba.²⁵ Both groups assigned ICD-9-CM codes to the conditions in the index, although the Dartmouth-Manitoba adaptation included some related diagnoses that were not in Charlson's original specification of 19 conditions. Both groups further demonstrated the ability of a claims-based Charlson index to predict mortality as well as other patient outcomes, such as postoperative complications, length of hospital stay, hospital charges, and discharge to nursing home care. Ghali et al²⁶ used hospital discharge data for patients who underwent coronary artery bypass graft surgery, and Cleves et al²⁷ used hospital claims data for patients with six different medical and surgical conditions to compare the performance of the Deyo and Dartmouth-Manitoba adaptations; these investigators found the two adaptations to be essentially equivalent.

The Kaplan and Feinstein index¹⁹ was developed nearly 30 years ago as a means of assessing comorbid conditions in patients with diabetes. It

TABLE 1. Established Indices or Algorithms for Measuring Comorbidity Using Claims Data

Index or Algorithm	Description	How it is Applied	Study Cohort and Prognostic Endpoints
Charlson Index ¹⁸	Index of 19 conditions, each assigned a weight according to its influence on mortality.	Each patient receives an integer score; a higher score indicates a greater burden of comorbid disease.	1-year all-cause mortality in 559 hospitalized medical service patients; 10-year non-breast cancer mortality in 685 breast cancer patients.
Kaplan and Feinstein ¹⁹	Score ranging from 0 to 3, with 0 indicating no comorbidity and 3 representing severe comorbidity.	Each patient is evaluated for the presence and severity of 12 conditions. The most severe comorbid disease determines the patient's comorbidity score.	5-year survival in 188 patients treated for diabetes at a Veteran's Administration hospital.
Satariano and Ragland ²⁰	Index of 7 conditions, unweighted, that were predictive of mortality.	Each patient receives an integer score; a higher score indicates a greater burden of comorbid disease.	3-year survival in 936 women diagnosed with breast cancer.
Count of No. of Unique Diagnosis Codes ²¹	Score ranging from 0 to 5 that reflects comorbidity as well as health status, potential resource use, and complications.	Each patient receives a score comprised of the number of diagnoses listed on the hospital claim; up to 5 diagnoses were possible.	Length of stay and 30-day mortality in 249,744 Medicare patients who underwent total knee replacement surgery.
Chronic Disease Score ²²	Integer score that reflects the number of prescription drugs taken for chronic disease and the extent to which the drug regimen is targeted toward life-threatening or progressive conditions.	For each patient, use of selected prescription drugs are assessed over a defined time period. Each patient receives an integer score, with higher scores indicative of a greater chronic disease burden.	Hospitalization and mortality during the year following evaluation of prescription drug use in 122,911 Group Health Cooperative of Puget Sound members.

has been used in a number of studies, although only one of these involved analysis of claims data.¹⁷ Newschaffer et al²⁸ adapted Kaplan and Feinstein's algorithm for use with Medicare inpatient data, and demonstrated its utility in predicting long-term survival in breast cancer patients. The Satariano and Ragland index,²⁰ comprising seven unweighted comorbid conditions, was developed in a cohort of breast cancer patients, is more recent, and also has been adapted for and used in only one study involving claims data.²⁸

Melfi et al²¹ showed in an analysis of Medicare claims data that a simple count of the number of unique diagnosis codes listed on the discharge summary was predictive of hospital length of stay and 30-day mortality. Rochon et al²⁹ obtained similar results, although their work was based on record abstraction rather than on claims data. While appealing in its simplicity, an important drawback to using a count of the number of unique diagnosis codes as a comorbidity measure is that the diagnoses listed in the claims database

or on the discharge abstract may represent complications of treatment or other manifestations, rather than true comorbid conditions.

The Chronic Disease Score²² was developed as a pharmacy-based measure of chronic illness burden, and may have utility in studies in which the claims database includes data on prescription drug use. Although at present Medicare claims do not contain prescription drug information, the addition of a prescription drug benefit to the Medicare program might make the Chronic Disease Score a viable option as a comorbidity measure for researchers working with Medicare or SEER-Medicare data.

To date, only two studies have directly compared the performance of multiple claims-based comorbidity or related measures.^{21,28} Melfi et al²¹ assessed the ability of three indices—the Charlson comorbidity index, the Relative Intensity Score (a measure of resource use), and a count of the number of unique diagnosis codes listed on the hospital claim—to predict length of stay and 30-day mortality in a large cohort of elderly Medicare patients. Although inclusion of each measure improved model fit, the Charlson index performed least well and the number of diagnoses best. Newschaffer et al²⁸ compared the ability of the Charlson,¹⁸ Kaplan and Feinstein,¹⁹ and Satariano and Ragland²⁰ comorbidity indices to predict 5-year survival in a cohort of breast cancer patients, and found the Charlson index to be the strongest predictor of long-term survival. These analyses lend support to the contention that the Charlson index performs best when long-term mortality is the outcome of interest.^{17,27}

Measurement of Comorbidity in Studies Using SEER-Medicare

As of February 2001, 15 published studies using the SEER-Medicare database have included a measure of comorbidity as a determinant of patterns of cancer treatment and/or outcomes.^{10,11,30-42} Fourteen of these studies^{10,11,30-31,33-42} have used some form of the Charlson index, eleven with the Deyo adaptation, and three with the Dartmouth-Manitoba adaptation.

While the use of the Charlson index in SEER-Medicare studies has been near universal, there has been less consistency among these projects in defining the period during which information about a patient's underlying health status is ob-

tained. Most SEER-Medicare investigators have reviewed inpatient claims over a 1- or 2-year period before the cancer was diagnosed to identify any prior hospitalizations that may report preexisting conditions. Looking retrospectively over a long period increases the number of claims from which diagnoses representing comorbid conditions may be extracted. This advantage is particularly beneficial for studies that use only inpatient claims to capture comorbidities, as only a small proportion of the Medicare population is hospitalized in a given year.⁴³ For persons with no hospital records, a comorbidity score of zero typically is assigned. Some investigators have dealt with limited comorbidity information by including concomitant diagnoses reported if the patient is hospitalized for cancer-directed surgery around the time of diagnosis, a strategy that greatly increases the number of persons for whom there is information about comorbidities. This approach should be viewed with caution, though, as it is not possible to distinguish conditions that were present before hospitalization (ie, before the cancer diagnosis) from complications that occurred during hospitalization.

An important caution about using claims that precede the cancer diagnosis to assess preexisting conditions is that for beneficiaries just becoming eligible for Medicare at age 65 years, there are no earlier years of data from which to obtain comorbidity information. Some researchers have dealt with this issue by limiting their sample to persons diagnosed at 66 or 67 years and older, depending on how many years of prediagnostic data are being reviewed. Other investigators have assigned a comorbidity score of zero to the newly eligible beneficiaries.

Of the studies cited in this review, all but one³⁴ have restricted determination of a patient's comorbidity status to a review of Medicare inpatient claims. This approach has predominated because, for many years, the inpatient claims were the only Medicare files that contained diagnoses. It is important to note that the Medicare physician and outpatient files now include diagnoses, which potentially increases the number of data sources that investigators might use to identify comorbidities. However, the accuracy of the diagnoses recorded in Medicare physician and outpatient files has received little study. Use of physician and outpatient claims to measure comorbidities is discussed further in the section on efforts to improve measurement of comorbidity using claims data.

Further Caveats About Claims Data

Claim databases such as SEER-Medicare are attractive to health services and outcomes researchers because they can provide large, population-based samples at relatively low costs. Investigators using these databases to measure comorbidity must be aware of the limitations of the data source, however. Claims databases typically are constructed for administrative rather than research purposes. For example, under Medicare's prospective payment system, the assignment and sequencing of diagnosis and procedure codes to describe the patient's hospital stay is accomplished primarily to obtain reimbursement rather than for statistical record-keeping purposes or to facilitate research studies. Also, because of limited information in the ICD-9 CM coding system, the diagnosis code appearing in claims may reflect that the person had a condition, without precise clinical detail. For example, while it may be possible to determine from ICD-9-CM codes in Medicare claims that a person has a diagnosis of diabetes, it may not be possible to know whether it is Type I or Type II diabetes or if there are diabetic complications. In addition, the assignment of specific ICD-9-CM codes may be open to interpretation by clinicians and medical records coders, resulting in variability in coding practices across clinicians, coders, or hospitals.¹

Studies of Medicare hospital claims that used data from the 1980s, after the Medicare prospective payment system was implemented, have documented that ICD-9-CM coding errors and coding accuracy vary by condition.⁴⁴⁻⁴⁷ A study using more recent data compared the principal ICD-9-CM diagnosis code on the hospital claim with the hospital medical record; there was 84.7% agreement at the five-digit level and 91.2% agreement at the three-digit level.⁴⁸ A study of the accuracy of the diagnosis on Medicare physician claims found high agreement between the physician claim and the office medical record for most diagnoses. The authors concluded that if the diagnosis appears on the claim, it is probably true, although they expressed concern about the completeness (eg, sensitivity) of the Medicare data.⁴⁹ The sensitivity of hospital claims data also varies by condition.^{25,45,48,52} This variability may result from the limited number of fields that are available for recording diagnoses and procedures, with conditions such as complications that provide more lucrative reimbursement receiving priority over

comorbid conditions. Validation studies comparing administrative claims and medical records as sources of comorbidity data generally have shown that comorbid conditions are underascertained in claims databases.⁵⁰⁻⁵³ This underascertainment has resulted in lower risk ratios than those obtained when the medical record is the comorbidity data source,^{4,25,50-51} although in at least two studies^{28,53} the relative mortality risk estimates associated with a one-level increase in comorbidity were nearly identical for claims-based and medical records-based Charlson scores.

Efforts to Improve the Measurement of Comorbidity Using Claims Data

Recognizing limitations in the data source and methodology, several groups of investigators within the past few years have attempted to develop improved measures of comorbidity for use in claims-based research. These efforts can be summarized into three categories: (1) augmentation of claims data with medical record, patient self-report, or health services utilization data; (2) incorporation of claims data from sources other than inpatient databases; and (3) exploration of alternative conditions, indices, or ways of grouping conditions.

Augmentation of Claims Data with Medical Record, Patient Self-Report, or Health Services Utilization Data. Because comorbid conditions tend to be underascertained in inpatient claims databases,^{44,50-51} some investigators have attempted to augment claims-based comorbidity measures with health services utilization data or comorbidity scores derived from medical records or patient self-reports. For example, Newschaffer et al²⁸ computed the Charlson,¹⁸ Kaplan and Feinstein,¹⁹ and Satariano and Ragland²⁰ comorbidity indices from both claims and medical record data for a cohort of breast cancer patients. Agreement between claims- and medical records-based comorbidity indices, as measured by the κ statistic, was modest at best, although both claims- and medical records-based comorbidity indices were predictive of mortality. Furthermore, the predictive ability of the mortality models was greatest when both the claims- and medical records-based Charlson indices were incorporated into the models. Similarly, Zhang et al⁵⁴ showed that the explanatory power of a claims-based Charlson index could be enhanced by including a comorbidity

measure derived from conditions self-reported by patients in the Medicare Current Beneficiary Survey. Finally, in an analysis of 1-year survival in breast cancer patients, Wang et al⁵⁵ showed that model performance improved when a measure of health care utilization (eg, number of prior hospital days) was incorporated along with the Charlson index.

Incorporation of Claims Data from Sources Other Than Inpatient Claims. Many claims-based studies that have used the Charlson comorbidity index have included only data from the inpatient setting.^{10,35,56} The trend toward delivery of health care services exclusively in outpatient settings may result in the omission of important comorbidities when only inpatient claims databases are considered.⁵⁷ Three groups of investigators have examined whether inclusion of claims from sources other than inpatient databases enhances measurement of comorbidity. Zhang et al⁵⁴ demonstrated improvements in the prediction of 2-year mortality in a sample of 1,387 Medicare patients by incorporating the Charlson comorbidity index derived from 1 year of outpatient and auxiliary claims, in addition to the Charlson comorbidity index derived from 2 years of inpatient claims. Klabunde et al⁵⁸ showed that a Charlson-based comorbidity measure derived from Medicare physician/supplier claims was predictive of 2-year mortality and receipt of less aggressive therapy in two large cohorts of cancer patients, in the presence of a separate Charlson-based comorbidity measure derived from inpatient claims. In contrast, Wang et al⁵⁵ concluded from an analysis involving a cohort of New Jersey women that expanding the sources of claims diagnoses to incorporate outpatient as well as inpatient care and assessing periods before an index hospitalization did not substantially improve the explanatory power of the Charlson index. However, this group combined comorbidities derived from inpatient and outpatient claims into a single comorbidity measure rather than establishing separate inpatient and outpatient comorbidity indices, as was done by Zhang et al⁵⁴ and Klabunde et al.⁵⁸

Exploration of Alternative Conditions, Indices, or Ways of Grouping Conditions. Most of the efforts to improve the measurement of comorbidity using claims databases that have been discussed so far have involved enhancements to the Charlson index. However, several limitations to the Charlson index must be recognized. The index comprises a finite number of chronic conditions

that were found to be prognostically important in a small sample of patients hospitalized at a single institution. Although most investigators applying Charlson's algorithm use the condition weights developed by Charlson, several groups have obtained different condition weights when replicating the algorithm with study-specific data.^{26-27,53,58-59} Furthermore, some evidence suggests that condition weights may not be necessary, and that a straightforward count of comorbid conditions may be equivalent to a weighted index that reflects the varying abilities of individual comorbid conditions to influence mortality.⁵⁵

Acknowledging the limited number of conditions represented in the Charlson index, Elixhauser et al⁶⁰ used a large statewide hospital discharge dataset to identify a comprehensive set of comorbid conditions that could be used in studies with a variety of endpoints. Seventeen different comorbid conditions were predictive of length of stay, hospital charges, and in-hospital mortality. Five of the original Charlson conditions (old myocardial infarction, cerebrovascular disease, dementia, renal disease, and leukemia) were dropped from this specification of important comorbidities, and three new conditions were added: coagulopathy, weight loss, and fluid and electrolyte disorders. These investigators recommended incorporating the conditions into models as separate measures rather than combining them into a summary comorbidity index, because their work showed that the conditions have significantly varying influences depending on the patient group and outcome of interest.

Some investigators have shown the utility of developing comorbidity indices that are tailored to a specific patient group. For example, Normand et al⁶¹ used expert opinion and logistic regression modeling to identify 21 comorbid conditions, and summarized these conditions in an index measure that proved to be a strong predictor of 2-year mortality in a cohort of patients hospitalized for acute myocardial infarction. Similarly, Polanczyk et al⁶² used logistic regression to create a comorbidity index consisting of six of the Charlson conditions plus four additional conditions (hyponatremia, other hydroelectric disturbances, ventricular arrhythmias, and hypotension/shock); the new index was a stronger predictor of in-hospital mortality than the Charlson index in patients hospitalized with congestive heart failure. Based on a literature review, Fleming et al⁶³ expanded Charlson's classification to include morbid obesity, major blood

vessel disease, immune disorders, and autoimmune diseases, and created a comprehensive prognostic index that predicted 1-year mortality in breast cancer patients.

Finally, the use of specialized software to identify and group conditions is an emerging and potentially promising area of investigation. Two such systems, the Adjusted Diagnostic Groups (ADG) case-mix system⁶⁴ and the Hierarchical Condition Categories (HCC) system,⁴³ use the diagnostic codes available in claims data to assign measures of health status—a broader construct than comorbidity—to individual patients. Both systems were designed to predict health care expenditures rather than mortality or type of therapy received. The ADGs were predictive of the number of ambulatory physician visits,⁶⁵ and have been used as a case-mix adjuster for practice patterns analyses in a large HMO.⁶⁶ To date, no studies have been published that compare the performance of ADGs or HCCs with the Charlson index as measures of underlying patient health status. Further work is needed to establish whether case-mix systems, initially developed to predict expenditures, can be adapted for use in assessing illness burden.

Conclusion

This overview described five established indices or algorithms that have been used to measure comorbidity in health services and outcomes research studies employing administrative or claims data: (1) the Charlson index, (2) the Kaplan and Feinstein index, (3) the Satariano and Ragland index, (4) a count of the number of unique diagnosis codes listed in claims, and (5) the Chronic Disease Score. Of these, the Charlson index is by far the most widely used, in studies involving Medicare claims as well as in other types of administrative databases. However, recognition of the limitations of the Charlson index and of ascertaining comorbid conditions from claims databases has prompted several investigators to explore ways of improving the measurement of comorbidity using claims data. These efforts have primarily focused on the Charlson index, and have entailed augmentation of claims data with medical record, patient self-report, or health services utilization data, or incorporation of claims data from sources other than inpatient claims.

Many opportunities for research in claims-based comorbidity assessment remain. For example, few studies have directly compared the performance of multiple claims-based comorbidity measures. In addition, there are established comorbidity measures such as the Index of Coexistent Disease⁶⁷ or the Cumulative Illness Rating Scale⁶⁸ that to date have not been adapted for use with administrative databases. Moreover, recent studies have shown the potential importance of several comorbid conditions that are not represented in the Charlson index, as well as the utility of developing comorbidity indices tailored to specific groups, such as patients with myocardial infarction, congestive heart failure, or breast cancer. Developing disease- or cohort-specific comorbidity measures may not be feasible or practical for many health services and outcomes researchers, however. Evaluation of whether measures based on health services utilization are biased due to variation in physician practice patterns or patient care-seeking behavior also has not received much attention.²² Finally, the use of specialized case-mix software such as the ADG and HCC systems, or the use of methodologies such as item response theory^{69,70} to measure underlying patient health status, are emerging areas of investigation that have not yet been compared with standard comorbidity algorithms such as the Charlson index but may afford relatively easy-to-apply and comprehensive approaches to health status measurement for investigators working with large, complex administrative databases.

In conclusion, the assessment of comorbidity using claims data is a complex and evolving area of investigation. There are multiple ways of measuring comorbidity using claims databases such as SEER-Medicare. Investigators should choose an approach based on the data they have available and the research question at hand, while remaining cognizant of the limitations of claims databases. No single comorbidity measure will completely account for differences in patients' underlying health status. Further research is needed to evaluate and improve comorbidity measurement with claims data.

References

1. Iezzoni LI. Risk Adjustment for Measuring Health Care Outcomes, 2nd ed. Chicago: Health Administration Press, 1997.

2. **Feinstein AR.** The pre-therapeutic classification of co-morbidity in chronic disease. *J Chron Dis* 1970;23:455–468.
3. **Van den Akker M, Buntinx F, Metsemakers JFM, et al.** Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent disease. *J Clin Epidemiol* 1998;51:367–375.
4. **Shwartz M, Iezzoni LI, Moskowitz MA, et al.** The importance of comorbidities in explaining differences in patient costs. *Med Care* 1996;34:767–782.
5. **Iezzoni LI, Heeren T, Foley SM, et al.** Chronic conditions and risk of in-hospital death. *Health Serv Res* 1994;29:435–460.
6. **Librero J, Peiro S, Ordinana R.** Chronic comorbidity and outcomes of hospital care: length of stay, mortality, and readmission at 30 and 365 days. *J Clin Epidemiol* 1999;52:171–179.
7. **Ogle KS, Swanson GM, Woods N, et al.** Cancer and comorbidity: redefining chronic disease. *Cancer* 2000;88:653–663.
8. **Greenfield S, Aronow HU, Elashoff RM, et al.** Flaws in mortality data: the hazards of ignoring comorbid disease. *JAMA* 1988;260:2253–2255.
9. **Newschaffer CJ, Penberthy L, Desch CE, et al.** The effect of age and comorbidity in the treatment of elderly women with nonmetastatic breast cancer. *Arch Intern Med* 1996;156:85–90.
10. **Ballard-Barbash R, Potosky AL, Harlan LC, et al.** Factors associated with surgical and radiation therapy for early stage breast cancer in older women. *J Natl Cancer Inst* 1996;88:716–726.
11. **Klabunde CN, Potosky AL, Harlan LC, et al.** Trends and black/white differences in treatment for nonmetastatic prostate cancer. *Med Care* 1998;36:1337–1348.
12. **Yancik R, Wesley MN, Ries LAG, et al.** Comorbidity and age as predictors of risk for early mortality of male and female colon carcinoma patients. *Cancer* 1998;82:2123–2134.
13. **West DW, Satariano WA, Ragland DR, et al.** Comorbidity and breast cancer survival: a comparison between black and white women. *Ann Epidemiol* 1996;6:413–419.
14. **Ribeiro K, Kowalski LP, Latorre M.** Impact of comorbidity, symptoms, and patients' characteristics on the prognosis of oral carcinomas. *Arch Otolaryngol Head Neck Surg* 2000;126:1079–1085.
15. **Practice Management Information Corporation.** International Classification of Diseases 9th Revision Clinical Modification, 5th ed. Los Angeles, CA: Practice Management Information Corporation, 1996.
16. **American Medical Association.** Physician's Current Procedural Terminology. Chicago, IL: American Medical Association, 1996:1995.
17. **Extermann M.** Measurement and impact of comorbidity in older cancer patients. *Crit Rev Oncol Hematol* 2000;35:181–200.
18. **Charlson ME, Pompei P, Ales KL, et al.** A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis* 1987;40:373–383.
19. **Kaplan MH, Feinstein AR.** The importance of classifying initial co-morbidity in evaluating the outcome of diabetes mellitus. *J Chron Dis* 1974;27:387–404.
20. **Satariano WA, Ragland DR.** The effect of comorbidity on 3-year survival of women with primary breast cancer. *Ann Intern Med* 1994;120:104–110.
21. **Melfi C, Holleman E, Arthur D, Katz B.** Selecting a patient characteristics index for the prediction of medical outcomes using administrative claims data. *J Clin Epidemiol* 1995;48:917–926.
22. **Von Korff M, Wagner EH, Saunders K.** A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 1992;45:197–203.
23. **Schneeweiss S, Maclure M.** Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol* 2000;29:891–898.
24. **Deyo RA, Cherkin DC, Cio IMA.** Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613–619.
25. **Romano PS, Roos LL, Luft HS, et al.** A comparison of administrative versus clinical data: coronary artery bypass surgery as an example. *J Clin Epidemiol* 1994;47:249–260.
26. **Ghali WA, Hall RE, Rosen AK, et al.** Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data. *J Clin Epidemiol* 1996;49:273–278.
27. **Cleves MA, Sanchez N, Draheim M.** Evaluation of two competing methods for calculating Charlson's comorbidity index when analyzing short-term mortality using administrative data. *J Clin Epidemiol* 1997;50:903–908.
28. **Newschaffer CJ, Bush TL, Penberthy LT.** Comorbidity measurement in elderly female breast cancer patients with administrative and medical records data. *J Clin Epidemiol* 1997;50:725–733.
29. **Rochon PA, Katz JN, Morrow LA, et al.** Comorbid illness is associated with survival and length of hospital stay in patients with chronic disability: a prospective comparison of three comorbidity indices. *Med Care* 1996;34:1093–1101.

30. Bach PB, Cramer LD, Warren JL, et al. Racial differences in the treatment of early-stage lung cancer. *N Engl J Med* 1999;341:1198–1208.
31. Begg C, Cramer LD, Hoskins WJ, et al. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA* 1998;280:1747–1751.
32. Cooper GS, Yuan Z, Chak A, et al. Geographic and patient variation among Medicare beneficiaries in the use of follow-up testing after surgery for nonmetastatic colorectal carcinoma. *Cancer* 1999;85:2124–2131.
33. Deleyiannis FW, Weymuller EA, Garcia I, et al. Geographic variation in the utilization of esophagoscopy and bronchoscopy in head and neck cancer. *Arch Otolaryngol Head Neck Surg* 1997;123:1203–1210.
34. Earle CC, Venditti LN, Neumann PJ, et al. Who gets chemotherapy for metastatic lung cancer? *Chest* 2000;117:1239–1246.
35. Lu-Yao GL, Potosky AL, Albertsen PC, et al. Follow up treatments after radical prostatectomy for prostate cancer - a population-based study. *J Natl Cancer Inst* 1996;88(3/4):166–172.
36. McCarthy EP, Burns RB, Coughlin SS, et al. Mammography use helps to explain differences in breast cancer stage at diagnosis between older black and white women. *Ann Intern Med* 1998;128:9:729–736.
37. Merrill RM, Brown ML, Potosky AL, et al. Colorectal cancer treatment and survival in Medicare HMO and Fee-for-Service settings. *Med Care Res Rev* 1999;56:177–196.
38. Potosky AL, Merrill RM, Riley GF, et al. Prostate cancer treatment and 10-year survival among group/staff HMO and FFS Medicare patients. *Health Serv Res* 1999;34:525–546.
39. Potosky AL, Merrill RM, Riley GF, et al. Breast cancer survival and treatment in HMO and fee-for-service settings. *J Natl Cancer Inst* 1997;89:1683–1691.
40. Schapira MS, McAuliffe TL, Nattinger AB. Underutilization of mammography in older breast cancer survivors. *Med Care* 2000;38(3):281–289.
41. Schrag D, Cramer LD, Bach PB, et al. Influence of hospital procedure volume on outcomes following surgery for colon cancer. *JAMA* 2000;284:3028–3035.
42. Warren JL, Riley GF, Potosky AL, et al. Trends and outcomes related to outpatient mastectomy in elderly women. *J Natl Cancer Inst* 1998;90:833–840.
43. Ellis RP, Pope GS, Iezzoni LI, et al. Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financ Rev* 1996;17:101–128.
44. Green J, Wintfeld N. How accurate are hospital discharge data for evaluating effectiveness of care? *Med Care* 1993;31:719–731.
45. Fisher ES, Whaley FS, Krushar WM, 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.
46. Hsia DC, Ahern CA, Ritchie BP, et al. Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988. *JAMA* 1992;268:896–899.
47. Iezzoni LI, Foley SM, Daley J, et al. Comorbidities, complications, and coding bias: does the number of diagnosis codes matter in predicting in-hospital mortality? *JAMA* 1992;267:2197–2203.
48. Romano PS, Luft HS. Getting the most out of messy data: problems and approaches for dealing with large administrative datasets. In: Grady ML, Schwarz M, eds. *Medical Effectiveness Research Data Methods*. Rockville, MD: Agency for Health Care Policy and Research, 1992.
49. Fowles JB, Lawthers AG, Weiner JP, et al. Agreement between physicians' office records and Medicare Part B claims data. *Health Care Financ Rev* 1995;16:189–199.
50. Malenka DJ, McLerran D, Roos N, et al. Using administrative data to describe casemix: a comparison with the medical record. *J Clin Epidemiol* 1994;47:1027–1032.
51. Kieszak SM, Flanders WD, Kosinski AS, et al. A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. *J Clin Epidemiol* 1999;52:137–142.
52. Hawker GA, Coyte PC, Wright JG, et al. Accuracy of administrative data for assessing outcomes after knee replacement surgery. *J Clin Epidemiol* 1997;50:265–273.
53. van Doorn C, Bogardus ST, Williams CS, et al. Risk adjustment for older hospitalized persons: a comparison of two methods of data collection for the Charlson index. *J Clin Epidemiol* 2001;54:694–701.
54. Zhang JX, Iwashyna TJ, Christakis NA. The performance of different look-back periods and sources of information for Charlson comorbidity adjustment in Medicare claims. *Med Care* 1999;37:1128–1139.
55. Wang PS, Walker A, Tsuang M, et al. Strategies for improving comorbidity measures based on Medicare and Medicaid claims data. *J Clin Epidemiol* 2000;53:571–578.
56. Desch C, Penberthy L, Newschaffer CJ, et al. Factors that determine the treatment for local and regional prostate cancer. *Med Care* 1996;34:152–162.
57. Silliman RA. Breast cancer care in old age: where do we go from here? *J Natl Cancer Inst* 1996;88:701–703.
58. Klabunde CN, Potosky AL, Legler JM, et al. Development of a comorbidity index using physician claims data. *J Clin Epidemiol* 2000;53:1258–1267.

59. **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.
60. **Elixhauser A, Steiner C, Harris DR, et al.** Comorbidity measures for use with administrative data. *Med Care* 1998;36:8-27.
61. **Normand SLT, Morris CN, Fung KS, et al.** Development and validation of a claims based index for adjusting for risk of mortality: the case of acute myocardial infarction. *J Clin Epidemiol* 1995;48:229-243.
62. **Polanczyk CA, Rohde LEP, Philbin EA, et al.** A new casemix adjustment index for hospital mortality among patients with congestive heart failure. *Med Care* 1998;36:1489-1499.
63. **Fleming ST, Rastogi A, Dmitrienko A, et al.** A comprehensive prognostic index to predict survival based on multiple comorbidities: a focus on breast cancer. *Med Care* 1999;37:601-614.
64. **Weiner JP, Dobson A, Maxwell SL, et al.** Risk-adjusted Medicare capitation rates using ambulatory and inpatient diagnoses. *Health Care Financing Review* 1996;77-99.
65. **Orueta JF, Lopez-De-Munain J, Baez K, et al.** Application of the Ambulatory Care Groups in the primary care of a European national health care system: does it work? *Med Care* 1999;37:238-248.
66. **Selby JV, Grunbach K, Quesenberry CP, et al.** Differences in resource use and costs of primary care in a large HMO according to physician specialty. *Health Serv Res* 1999;34:503-518.
67. **Greenfield S, Blanco DM, Elashoff RM, et al.** Patterns of care related to age of breast cancer patients. *JAMA* 1987;257:2766-2770.
68. **Linn BS, Linn MW, Gurel L.** Cumulative illness rating scale. *J Am Geriatr Soc* 1968;16:622-666.
69. **Embretson SE, Reise SP.** Item Response Theory for Psychologists. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
70. **Krueger RF, Finger MS.** Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychological Assess* 2001;13:140-11.