



Published in final edited form as:

Med Care. 2006 August ; 44(8): . doi:10.1097/01.mlr.0000223475.70440.07.

In Search of the Perfect Comorbidity Measure for Use With Administrative Claims Data:

Does It Exist?

Laura-Mae Baldwin, MD, MPH*, Carrie N. Klabunde, PhD[†], Pam Green, PhD, MPH*, William Barlow, PhD[‡], and George Wright, PhD*

*Department of Family Medicine University of Washington, Seattle, Washington [†]Department of Biostatistics University of Washington, Seattle, Washington [‡]Division of Cancer Control and Population Sciences, Applied Research Program, National Cancer Institute, Bethesda, Maryland.

Abstract

Background—Numerous measures of comorbidity have been developed for health services research with administrative claims.

Objective—We sought to compare the performance of 4 claims-based comorbidity measures.

Research Design and Subjects—We undertook a retrospective cohort study of 5777 Medicare beneficiaries ages 66 and older with stage III colon cancer reported to the Surveillance, Epidemiology, and End Results Program between January 1, 1992 and December 31, 1996.

Measures—Comorbidity measures included Elixhauser's set of 30 condition indicators, Klabunde's outpatient and inpatient indices weighted for colorectal cancer patients, Diagnostic Cost Groups, and the Adjusted Clinical Group (ACG) System. Outcomes included receipt of adjuvant chemotherapy and 2 year noncancer mortality.

Results—For all measures, greater comorbidity significantly predicted lower receipt of chemotherapy and higher noncancer death. Nested logistic regression modeling suggests that using more claims sources to measure comorbidity generally improves the prediction of chemotherapy receipt and noncancer death, but depends on the measure type and outcome studied. All 4 comorbidity measures significantly improved the fit of baseline regression models for both chemotherapy receipt (baseline c-statistic 0.776; ranging from 0.779 after adding ACGs and Klabunde to 0.789 after Elixhauser) and noncancer death (baseline c-statistic 0.687; ranging from 0.717 after adding ACGs to 0.744 after Elixhauser).

Conclusions—Although some comorbidity measures demonstrate minor advantages over others, each is fairly robust in predicting both chemotherapy receipt and noncancer death. Investigators should choose among these measures based on their availability, comfort with the methodology, and outcomes of interest.

Copyright © 2006 by Lippincott Williams & Wilkins

Reprints: Laura-Mae Baldwin, MD, MPH, University of Washington, Department of Family Medicine, Box 354982, Seattle, WA 98195-4982. lmb@fammed.washington.edu..

The views expressed in this article are those of the authors and do not necessarily represent the views of the National Cancer Institute.

Preliminary results from this study were presented as a poster at the San Diego, California, AcademyHealth national meeting in June 2004.

Keywords

comorbidity; claims data; Medicare; risk adjustment

Administrative claims data are used frequently in health services research and policy to examine patient health status and the quality and cost of care. Comorbid illness has a significant impact on these and other outcomes, and several measures of comorbidity have been developed for use in studies based on administrative claims.^{1–8} Claims-based measures of comorbidity are of particular importance to cancer care researchers, who increasingly use population-based cancer registry data linked with administrative claims (eg, Surveillance, Epidemiology, and End Results [SEER]-Medicare files) to examine such issues as racial or geographic disparities in receipt of recommended treatments or the association between treatment and survival.^{9–12} Adjustment for comorbidity is essential in these observational studies because baseline differences in health status between groups (eg, racial and ethnic groups, those with and without treatment) may modulate differences found in study outcomes.

Health services researchers have a choice of comorbidity measures and must decide how best to apply them to their work. The measures themselves have been developed and used with claims from different places of services (eg, inpatient, outpatient) and time frames in relationship to the disease diagnosis (eg, prediagnosis, during treatment or diagnosis hospitalization). Some measures were developed to predict cost and modify payment systems,^{13–15} others to predict mortality and control for preexisting health status.^{1–8} A few studies have compared the performance of different measures in predicting mortality, generally concluding that measures incorporating more conditions and data sources are better predictors of mortality. Although some investigators have found no advantage to disease-specific indices, several have suggested creating condition- and/or outcome-specific comorbidity indices. Creating study-specific measures may not be possible, however, and investigators often apply measures in a manner at variance from how they were developed.^{10,12,16–18}

In this study, we identified 4 administrative claims-based measures of comorbidity that have been used frequently in the health services literature to adjust for baseline health status and compared their performance in predicting 2 outcomes of importance to cancer care researchers—receipt of recommended adjuvant chemotherapy and mortality—among stage III colon cancer patients. We chose to study stage III colon cancer patients for several reasons. First, colon cancer is generally a disease of the elderly, who are more likely to have comorbid illness. Second, a 1990 National Institutes of Health Consensus Panel on Colorectal Cancer recommended routine adjuvant chemotherapy for stage III colon cancer, yet the reported rate of adjuvant chemotherapy use among the elderly is only about 55%.¹⁹ Because comorbidity may contribute to this low treatment rate, identifying the optimal comorbidity measure is important for analyses examining receipt of recommended treatment. Third, colon cancer patients' level of comorbid illness could influence survival both through the influence on receipt of recommended treatment and through the association with noncancer death.

This work compares the performance of a wide variety of comorbidity measures in predicting treatment use and mortality, and examines the influence of using different data sources on measure performance. Because one of our measures was developed specifically for colorectal cancer patients, we were able to compare its performance to that of other measures derived from more generalized populations. We hypothesized that a measure that

combines multiple data sources with specificity to colorectal cancer patients would be the best predictor of our study outcomes.

METHODS

Database

This study used data from the National Cancer Institute's SEER cancer registries linked with Medicare claims for persons found in both files. The SEER program abstracts medical records to collect patient demographics, tumor characteristics, stage at diagnosis, and cause of death information for persons with cancer. Twelve SEER registries covering about 14% of the U.S. population were included.

Medicare data included vital status and enrollment status in Parts A and B Medicare and in Health Maintenance Organizations by month and year, extracted from the Centers for Medicare & Medicaid Services' denominator files. The Medicare Provider Analysis and Review (MedPAR) file includes all Part A short stay, long stay, and skilled nursing facility bills for each calendar year. Each MedPAR record includes up to 10 International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnoses. The Part B Physician/Supplier and Outpatient facility files consist of claims from physicians and other noninstitutional providers and institutional outpatient providers (eg, hospital outpatient departments). Each claim in the Physician/Supplier file allows the provider to record up to 4 ICD-9-CM diagnoses in addition to the diagnosis that accompanies each billed procedure. Each claim in the Outpatient facility file can include up to 10 ICD-9-CM diagnoses.

Study Population

From the SEER database, we identified 9894 Medicare beneficiaries ages 66 and older diagnosed with stage III colon cancer (SEER site codes 18.2–18.9, 19.9) between 1992 and 1996. We excluded patients with tumor histology other than adenocarcinoma ($n = 105$), simultaneous stage IV colorectal cancer identified in SEER records ($n = 11$), previous colorectal cancer ($n = 269$), and autopsy- or death certificate-based cancer diagnosis ($n = 3$). To optimize ascertainment of comorbid conditions, we excluded patients who lacked complete enrollment in both Parts A and B fee-for-service Medicare for 1 year prior to cancer diagnosis ($n = 2242$). To ensure optimal ascertainment of adjuvant chemotherapy, we excluded patients who were not alive and completely enrolled in both Parts A and B fee-for-service Medicare for 9 months following cancer diagnosis ($n = 1331$). We excluded patients without an index resection surgery claim within 6 months of colorectal cancer diagnosis ($n = 156$), as they may not have received definitive treatment. Our final study population included 5777 stage III colon cancer patients. Permission to conduct this study was granted by the University of Washington's Human Subjects Division.

Comorbidity Measures

Measure Description—We compared 4 previously published comorbidity measures developed for use with administrative data.^{2,5,7,13,15} The first is Klabunde's adaptation of the Charlson comorbidity index. The Charlson index² consists of 19 comorbid conditions weighted according to the degree to which they predicted mortality among an inpatient cohort, then summed to produce an index score. Klabunde adapted the Charlson index to create disease-specific comorbidity indices—one using the MedPAR claims and one using the Physician/Supplier and Outpatient facility claims.⁷ Each was constructed by applying the estimated coefficients derived from a Cox proportional hazards model with two-year noncancer mortality as the outcome to the 14 individual condition indicators and summing the weighted conditions. Klabunde prioritized inpatient diagnoses, constructing the

outpatient claims index using only those diagnoses not found in the inpatient claims. The present study applies a set of weights specifically developed for colorectal cancer patients (J. M. Legler, C. N. Klabunde, J. L. Warren, et al, unpublished data, January 2005).

The second comorbidity measure was developed by Elixhauser for use with large administrative databases to predict mortality, hospital charges, and hospital length of stay.⁵ Elixhauser used both Diagnosis-Related Groups (DRGs) and secondary ICD-9-CM diagnosis codes from inpatient claims to identify 30 unweighted comorbidity indicators that are entered as separate indicator (dummy) variables in a regression model. Software is available from the Healthcare Cost and Utilization Project to create these measures.²⁰

This study's third measure, DCGs, was developed originally to predict health care costs for Medicare payment models and has been adapted for risk adjustment purposes.²¹ This study implemented version 6.1 of DxCG's Risk Adjustment Software.²² The DCG models use inpatient and outpatient ICD-9-CM diagnosis codes, aggregating those that are clinically homogeneous into 781 DxGroups, then organizing them into 184 Condition Categories (CCs) that are similar both clinically and in level of resource use. Along with age and sex, hierarchies of Condition Categories are used to predict resource use either in the year concurrent with the diagnoses or in the following year (prospective models). We used the prospective DCG models in this study, as these are most likely to reflect individuals' chronic illness burden.

The fourth measure was adapted from the Adjusted Clinical Group (ACG) System, Version 5.0, a diagnosis-based, case-mix methodology that describes or predicts a population's past or future healthcare use and costs.¹⁵ The ACG System first assigns each patient's inpatient and outpatient diagnosis codes of all types to 1 of 32 Adjusted Diagnostic Groups (ADGs). The ADGs are subdivided into 8 "major" adult ADGs, which have very high expected resource use, and 24 "minor" adult ADGs. A combination of ADGs, age, and gender places individuals with similar morbidity and resource consumption into 1 of 93 discrete ACG categories.

Construction of Comorbidity Variables—This study was intended as an applied comparison of existing comorbidity measures. Thus, we developed comorbidity variables in the manner that they were originally defined, as this most closely replicates their use in published research. We applied 2 exclusions to the development of all of our measures to ensure consistent claims use. We excluded diagnosis codes from clinical laboratory, diagnostic imaging, and durable medical equipment claims in the Part B claims, as the codes associated with these claims may not be assigned by clinicians. To avoid "rule out" diagnoses and prevent overestimation of comorbidity when using Physician/Supplier or Outpatient facility claims, we required diagnosis codes on these claims to be listed in at least 2 different claims occurring more than 30 days apart. In addition, we excluded all diagnosis codes for colorectal and other intestinal cancers to avoid counting the cancer diagnosis as a comorbidity. For the variables describing comorbidity before the cancer diagnosis, we used claims in the 11 months prior to the month before diagnosis to avoid overestimating comorbidity related to the colon cancer diagnosis.

Both the DCG and ACG measures were designed to use a combination of inpatient and outpatient claims from a baseline time period. We used the combined diagnoses from the MedPAR, Physician/Supplier, and Outpatient facility claims to create these variables. The DCG measure used these data to assign individuals to 1 of 25 categories of predicted expenditures in the following year. Because of small numbers of patients in individual categories, we aggregated them into 10 distinct predicted expenditure categories and used them as individual, unweighted dummy variables in our analyses. The ACG System

identified individuals with 8 major and 24 minor Adjusted Diagnostic Groups based on expected resource use. Preliminary studies demonstrated a strong relationship between major ADGs and our study outcomes. We counted each individual's number of major ADGs before cancer diagnosis and created 4 unweighted ADG categories that were used as dummy variables in our analyses—0, 1, 2, 3 or more major ADGs.

The Klabunde measure creates 2 mutually exclusive comorbidity indices from the claims before cancer diagnosis, one reflecting inpatient diagnoses and the other outpatient diagnoses not found in inpatient claims. Klabunde found that adding an outpatient index contributed significantly to the prediction of recommended cancer treatment as well as 2-year mortality among breast and prostate cancer patients. The addition of an outpatient index ensures more complete representation of an individual's comorbid conditions, as only 16% of our study patients had inpatient hospitalizations in the baseline period prior to cancer diagnosis. Klabunde created mutually exclusive comorbidity variables because of the redundancy between inpatient and outpatient claims, and prioritized inpatient claims over outpatient claims a priori with the expectation that they would be stronger predictors. We maintained this strategy because we wished to identify the influence of adding claims from different settings and time periods to the prediction of our study outcomes. Building on these 2 indices, we created a third mutually exclusive comorbidity index representing conditions listed during the cancer resection hospitalization and not previously found in the inpatient and outpatient claims prior to the cancer diagnosis.

Elixhauser's set of 30 unweighted condition indicators were originally developed using inpatient data only. We adapted Elixhauser's measure to include diagnoses from outpatient claims data (ie, Physician/Supplier and Outpatient facility files) in a manner similar to that of Klabunde. We created 3 mutually exclusive sets of unweighted condition indicators using prior inpatient claims, prior outpatient claims, and cancer resection inpatient claims.

Outcome Measures

Outcome measures include receipt of adjuvant chemotherapy within 9 months of diagnosis, and death due to noncancer causes within 24 months of diagnosis. Patients who received chemotherapy had at least 1 claim for chemotherapy administration (Current Procedural Terminology codes 96408, 96410, 96412, 96414, 96520, 96530, 96545, 96549; ICD-9-CM codes 99.25, E0781, V58.1; Healthcare Common Procedure Coding System [HCPCS] codes J0640, J9190, Q0083, Q0084, Q0085). Cause of death was determined from the SEER variable Vital Status by Cause of Death, which uses state death certificates as a primary data source.²³ We emphasized noncancer death as an outcome because in analyses examining cancer care outcomes, comorbidity variables are used to control for the influence that noncancer conditions may have on the study outcomes. Patients classified as having died of cancer or died of unknown causes were excluded from the analyses with noncancer death as the outcome. Though we could find no studies validating colon cancer as the cause of death, a study validating cause of death from SEER data with hospital records for men with prostate cancer found 97% agreement between these 2 sources.²⁴ The sample for these analyses comprised 4003 stage III colon cancer patients (mortality cohort).

Analyses

We first described the sociodemographic characteristics (age, race, gender, marital status), environmental characteristics (SEER region, census tract-based race/age-specific median income, census tract-based race-specific percentage of high school graduates among 25 year and older persons) and vital status of our study population. We then calculated the proportion of the study population with comorbid conditions for each of our comorbidity measures. We used logistic regression to assess the contribution of each comorbidity

measure to prediction of chemotherapy receipt and noncancer death, after controlling for a set of baseline variables (age, race, gender, marital status, median income, percentage of high school graduates, and geographic region). We chose to control for these variables because they are commonly used in studies employing administrative data sources, and we wished to measure the performance of comorbidity variables independent of other factors that can influence our study outcomes. Likelihood ratio tests were used to determine whether adding sources of comorbidity data (eg, outpatient and inpatient versus inpatient only) improved the fit of our regression models. We report the Akaike Information Criterion (AIC) and the c-statistic²⁵ to evaluate the overall goodness of fit of logistic regression models with different comorbidity measures and claims sources. These measures allow descriptive comparison of non-nested models since hypothesis testing is restricted to nested models. The AIC accounts for the number of covariates in the model, a factor of importance for the Elixhauser measure, with 30 individual indicator variables. To assess the variability of the c-statistic, we performed bootstrap resampling with 1000 replications. The 95% confidence interval is percentile-based and uses the 2.5 and 97.5 percentiles of the distribution of the bootstrapped c-statistics.²⁶ This procedure does not guarantee that the observed statistic is inside the confidence interval. The percentile-based confidence interval can be biased and may be affected by the limited range of the c-statistics (0.50 to 1.00). It is given only to characterize the amount of variability in this measure.

RESULTS

Among the 5777 stage III colon cancer patients and the subset of 4003 individuals in the mortality cohort, nearly 3 quarters were 80 year olds or younger, more than 80% were white, and about half were married (Table 1). About 70% of the study population lived in census tracts in which over 3 quarters of the population had graduated from high school. Within 2 years, 39% of the overall study population had died.

Figure 1 displays the proportion of the study population with at least 1 comorbid condition using the 4 comorbidity measures and demonstrates how adding claims sources increases the identification of individuals as having comorbid disease. Using only prior inpatient claims, less than 20% of individuals were identified as having comorbid conditions in the Elixhauser and Klabunde measures. This proportion at least doubled when diagnoses from prior outpatient claims were added, and increased substantially again with the addition of cancer resection inpatient claims. Even using only prior inpatient and outpatient claims, the DCG measure identifies a large proportion of individuals as having comorbidity, since it includes all diagnosis codes. Overall, the Klabunde and major ADG measures identified fewer individuals with comorbidity, as they used fewer ICD-9-CM diagnoses.

In Table 2, we report frequencies for the 4 comorbidity measures and the degree to which these measures predict study outcomes. We report only those Elixhauser conditions significantly associated with at least one of our study outcomes. (A complete regression model is available from the authors upon request.) In general, individuals with comorbidity are less likely to receive chemotherapy, and more likely to die of a condition other than cancer within 2 years. A notable exception is that a prior diagnosis of hypertension in the Elixhauser measure increases rather than decreases an individual's likelihood of receiving chemotherapy.

Figure 2 reports the change in deviance ($-2 \log$ likelihood) from the baseline model to illustrate the improvement in goodness of fit as additional claims sources for identifying comorbidity are added. For both comorbidity measures, the addition of comorbid diagnoses from the prior inpatient claims to the baseline model significantly improves the fit of the regression models for both receipt of chemotherapy and non-cancer death. However, the

addition of comorbidity diagnoses from prior outpatient claims significantly improves the fit of both regression models only for noncancer death. For receipt of chemotherapy, the addition of prior outpatient claims diagnoses significantly improves model fit for the Elixhauser measure only. Addition of diagnoses from the cancer resection hospitalization improves the fit of the receipt of chemotherapy and noncancer death models for both measures.

To compare the 4 comorbidity measures, we used a standard model, including previous inpatient and previous outpatient claims. All 4 measures significantly improved the fit of the regression model beyond the baseline model as measured by a statistically significant change in the likelihood ratio (not shown). Figure 3 presents the c-statistics and AICs for the models examining receipt of chemotherapy and noncancer death as outcomes. The c-statistic suggests that the Elixhauser measure produces the best fitting model for both study outcomes. However, the AIC measure, which penalizes the model by the number of covariates used, shows the Elixhauser model has a poorer fit, making it difficult to conclude that any 1 model outperforms the others for either outcome. The small increase in c-statistic after adjusting for each of the 4 comorbidity measures suggests that comorbidity has a relatively minor influence on chemotherapy receipt after controlling for other sociodemographic and environmental variables.

We conducted several subanalyses to test the robustness of our study findings (not shown). First, because questions have been raised about the validity of the cause of death variable, we examined all-cause death as an outcome using the overall study sample. We found that, as for noncancer death, there was a significant association between each of the comorbidity measures and all-cause death, although the measures were more powerful predictors of noncancer death (ie, odds ratios farther from one) than all-cause death. Second, to ensure that our results were generalizable to a broader patient population, we included stage I–III colon cancer patients rather than only stage III colon cancer cases in the regression model using noncancer death as the outcome and obtained similar results. Each of the measures significantly improved the fit of both models. Third, to explore the degree to which the high performance of the Elixhauser measure was related to its use of a large number of individual indicator variables, we aggregated the individual Elixhauser condition indicators into a variable that counted the number of conditions. We divided them into 4 groups—0, 1, 2, 3 or more—and created indicator variables for each group. The c-statistics for the regression model using this Elixhauser condition count (final c-statistic 0.710 for noncancer death outcome; 0.775 for receipt of chemotherapy outcome) were lower than both those using the individual Elixhauser conditions (c-statistic 0.755, 0.788 respectively) and those using the Klabunde measure (c-statistic 0.731, 0.776, respectively).

DISCUSSION

Health services researchers using administrative claims data are faced with several methods and data sources for comorbidity adjustment and are often uncertain about which to choose. This study compared 4 comorbidity measures derived from outpatient and inpatient Medicare claims data and found that all were significant predictors of 2 very different outcomes—receipt of chemotherapy and noncancer death—in stage III colon cancer patients. Each measure significantly improved the fit of regression models, as measured by the change in log likelihood beyond a baseline model that included demographic and environmental variables. The contribution of comorbidity to the models predicting chemotherapy receipt was more minor, however, consistent with prior literature.⁷

No single comorbidity measure consistently outperformed the others in our regression models. Measures performing best using 1 criterion did not necessarily do so with another.

Therefore, investigators may wish to consider practical considerations in their measure choices. Overall, the Elixhauser measure demonstrated the greatest change in -2 log likelihood when adding different claim types to the model and the highest c -statistic when comparing the 4 measures. This is consistent with other studies comparing the performance of the Charlson and Elixhauser measures.^{27,28} When we aggregated the Elixhauser measure into a count of the number of conditions, its performance deteriorated substantially, suggesting that both the large number of conditions and their use as individual indicators contribute to its enhanced performance. If data from several mutually exclusive sources (eg, inpatient claims, outpatient claims) are included in a regression model, this 30-item comorbidity measure can quickly multiply to 60 or 90 individual variables and is only practical with a large administrative database. If these requirements are met, however, the Elixhauser measure has the advantage of being able to demonstrate an association between individual medical conditions and the study outcomes. In our study, for example, we identified the association between hypertension and an increased likelihood of chemotherapy receipt and a decreased likelihood of death, perhaps because it is a more minor health condition that ensures an ongoing relationship with a health provider or because hypertension was more likely to be coded among healthier patients without multiple, more severe diagnoses. Our finding is consistent with the results of 2 other studies that demonstrated lower in-hospital mortality rates among patients with hypertension.^{5,28}

The contribution of the different claim types varied by study outcome and comorbidity measure. When examining noncancer death, both the Elixhauser and Klabunde measures benefited from including prior outpatient, prior inpatient, and resection hospitalization claims data. For the Klabunde measure, the prior outpatient claims contributed to the model more significantly than the resection hospitalization claims. When examining receipt of adjuvant chemotherapy, all 3 claim types significantly improved the model fit for the Elixhauser measure, whereas use of the prior inpatient and resection hospitalization claims only improved the model fit for the Klabunde measure. These findings suggest that if using the Elixhauser measure, creating separate variables for each of the 3 claim types may be worthwhile. For the Klabunde measure, using all 3 claim types is important for examining the noncancer death outcome, while prior inpatient and cancer resection hospitalization claims could be used for the receipt of chemotherapy outcome.

Given the relatively equivalent performance of the comorbidity measures, investigators may wish to base their selection on theoretical and statistical considerations. For example, the Klabunde measure has been developed for use in studies of treatment and mortality in cancer patients, and different condition weights have been calculated for patients with different types of cancers.⁷ Klabunde's measure also sums estimated coefficients from multivariate Cox models, resulting in a scale that is consistent with the Cox proportional hazards and logistic regression models in which it is frequently applied. The ACGs and DCGs were initially developed to predict health care costs, making these measures preferable in models using cost as an outcome. Subsequent studies have demonstrated the utility of the DCGs in predicting mortality as well.²¹

This study is limited by its use of a narrow population group, stage III colon cancer patients who lived for at least 9 months after cancer diagnosis, and 2 relatively short-term outcome measures. Results may not be generalizable to other population groups or outcome measures, although our analysis demonstrating similar findings among stage I–III colon cancer patients is encouraging. There are limitations to the study's data source as well. Claims databases generally are constructed for administrative rather than research purposes, and the accuracy of the ICD-9-CM codes used to identify comorbidities in these databases may be variable. Moreover, comorbid conditions have been shown to be underascertained in claims when compared with the medical record as data sources.²⁹ In addition, there may be

some misclassification of acute complications as comorbid conditions (eg, electrolyte disorders), especially among inpatient admissions, resulting in overascertainment of conditions unrelated to the cancer diagnosis.

Despite these limitations, this analysis of the performance of 4 comorbidity measures in predicting chemotherapy receipt and noncancer mortality among stage III colon cancer patients adds to a growing literature evaluating the utility of comorbidity measures in predicting a variety of outcomes among different patient populations. While there are minor advantages of some measures over others, and of adding claims of different types and from different time periods, each of the measures is fairly robust in its ability to predict this study's outcomes. Thus, investigators can choose from among these comorbidity measures based on data availability and their study objectives, as well as access to and comfort with the methodology.

Acknowledgments

We thank Barbara Matthews for her expertise in designing and developing the study database and a number of its variables; Yong Cai, PhD, for his assistance in developing several of the study variables; Denise Lishner for her assistance in literature review; and Rachel Ballard-Barbash, PhD, and Martin Brown, PhD, from the National Cancer Institute for their review of the manuscript.

Supported by grant R01CA089544 from the National Cancer Institute, National Institutes of Health, Bethesda, Maryland.

REFERENCES

1. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol.* 1992; 45:613–619. [PubMed: 1607900]
2. 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. discussion 1081–1090. [PubMed: 8410092]
3. Normand SL, 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. [PubMed: 7869069]
4. 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. [PubMed: 8676173]
5. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. *Med Care.* 1998; 36:8–27. [PubMed: 9431328]
6. 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. [PubMed: 10386572]
7. Klabunde CN, Potosky AL, Legler JM, et al. Development of a comorbidity index using physician claims data. *J Clin Epidemiol.* 2000; 53:1258–1267. [PubMed: 11146273]
8. Fleming ST, Pearce KA, McDavid K, et al. The development and validation of a comorbidity index for prostate cancer among Black men. *J Clin Epidemiol.* 2003; 56:1064–1075. [PubMed: 14614997]
9. Steyerberg EW, Earle CC, Neville BA, et al. Racial differences in surgical evaluation, treatment, and outcome of locoregional esophageal cancer: a population-based analysis of elderly patients. *J Clin Oncol.* 2005; 23:510–517. [PubMed: 15659496]
10. Shavers VL, Brown ML, Potosky AL, et al. Race/ethnicity and the receipt of watchful waiting for the initial management of prostate cancer. *J Gen Intern Med.* 2004; 19:146–155. [PubMed: 15009794]
11. Hershman D, Jacobson JS, McBride R, et al. Effectiveness of platinum-based chemotherapy among elderly patients with advanced ovarian cancer. *Gynecol Oncol.* 2004; 94:540–549. [PubMed: 15297201]

12. Schrag D, Panageas KS, Riedel E, et al. Surgeon volume compared to hospital volume as a predictor of outcome following primary colon cancer resection. *J Surg Oncol*. 2003; 83:68–78. discussion 78–69. [PubMed: 12772198]
13. DxCG, Inc.. DxCG Risk Adjustment Software. Analytic Guide Release 6.1. Author; Boston: 2002.
14. DxCG, Inc.. DxCG Risk Adjustment Software. User's Guide Release 6.1. Author; Boston: 2002.
15. Weiner, JP.; Abrams, C., editors. The Johns Hopkins ACG Case-Mix System. Documentation & Application Manual. Johns Hopkins University; Baltimore, MD: 2001.
16. Piccirillo JF, Spitznagel EL Jr, Vermani N, et al. Comparison of comorbidity indices for patients with head and neck cancer. *Med Care*. 2004; 42:482–486. [PubMed: 15083109]
17. Parker JP, McCombs JS, Graddy EA. Can pharmacy data improve prediction of hospital outcomes? Comparisons with a diagnosis-based comorbidity measure. *Med Care*. 2003; 41:407–419. [PubMed: 12618644]
18. 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. [PubMed: 8526702]
19. Schrag D, Cramer LD, Bach PB, et al. Age and adjuvant chemotherapy use after surgery for stage III colon cancer. *J Natl Cancer Inst*. 2001; 93:850–857. [PubMed: 11390534]
20. Healthcare Cost and Utilization Project. [Accessed April 20, 2006] Comorbidity software. version 3.0 Available at: <http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp>.
21. Ash AS, Posner MA, Speckman J, et al. Using claims data to examine mortality trends following hospitalization for heart attack in Medicare. *Health Serv Res*. 2003; 38:1253–1262. [PubMed: 14596389]
22. DxCG Risk Adjustment Software. Release 6.1. DxCG, Inc.; Boston: 2002.
23. Bach PB, Guadagnoli E, Schrag D, et al. Patient demographic and socioeconomic characteristics in the SEER-Medicare database applications and limitations. *Med Care*. 2002; 40:IV-19–IV-25.
24. Penson DF, Albertsen PC, Nelson PS, et al. Determining cause of death in prostate cancer: are death certificates valid? *J Natl Cancer Inst*. 2001; 93:1822–1823. [PubMed: 11734600]
25. Hosmer, DW.; Lemeshow, S. Applied Logistic Regression. 2nd ed. Wiley; New York: 2000.
26. Efron, B. The Jackknife, the Bootstrap, and Other Resampling Plans. Society for Industrial and Applied Mathematics; Philadelphia: 1982.
27. Stukenborg GJ, Wagner DP, Connors AF Jr. Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. *Med Care*. 2001; 39:727–739. [PubMed: 11458137]
28. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. *Med Care*. 2004; 42:355–360. [PubMed: 15076812]
29. Klabunde CN, Warren JL, Legler JM. Assessing comorbidity using claims data: an overview. *Med Care*. 2002; 40:IV-26–35.

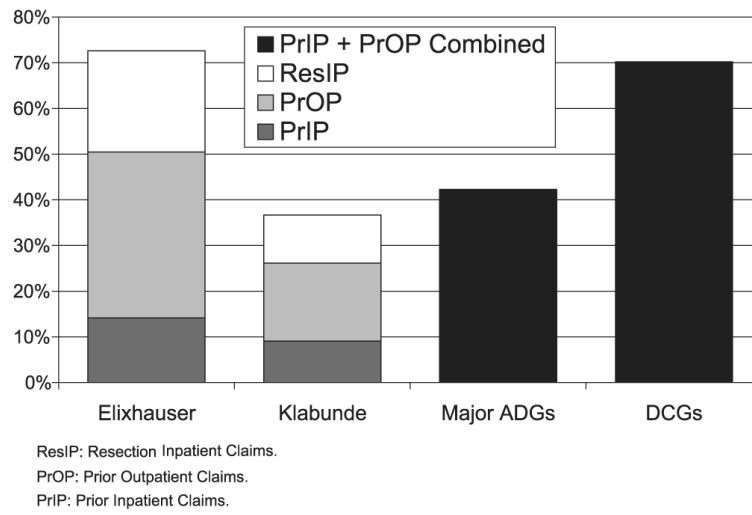


FIGURE 1. Percentage of patients with any comorbidity identified from different claims sources (n = 5777).

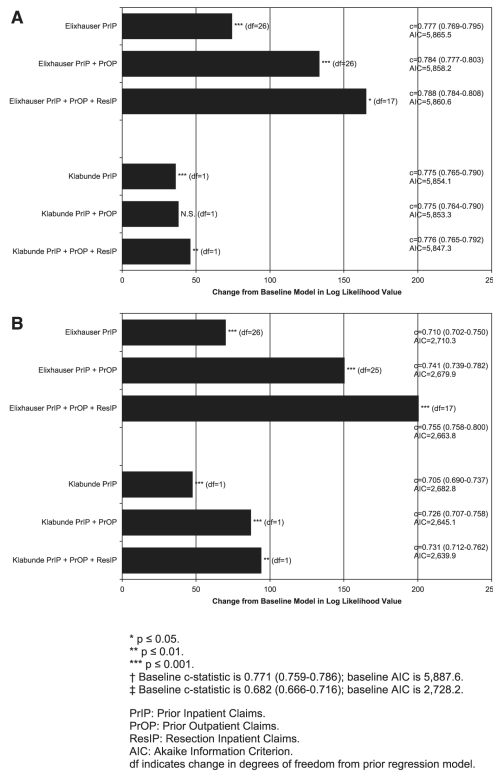


FIGURE 2. A, Influence of different claim sources on logistic regression model fit by comorbidity measure where outcome is receipt of chemotherapy.[†] B, influence of different claim sources on logistic regression model fit by comorbidity measure where outcome is non-cancer death within 24 months of diagnosis.[‡]

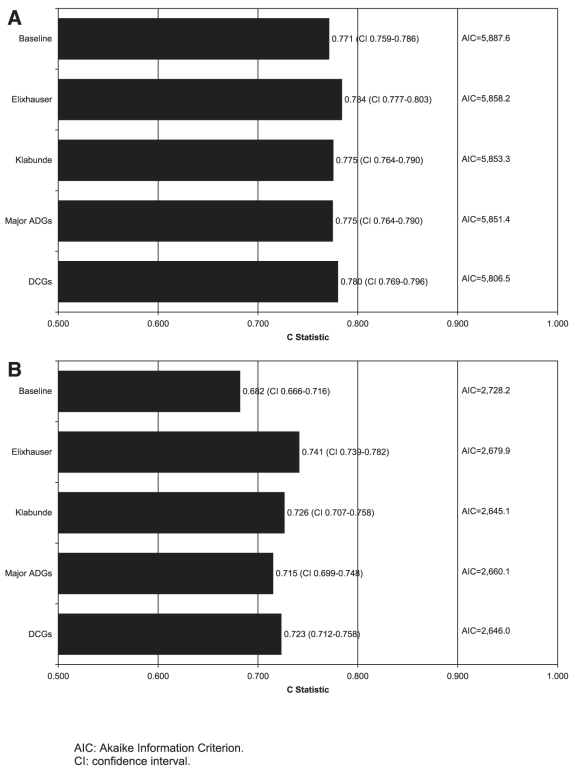


FIGURE 3. A, Performance of different comorbidity measures where outcome is receipt of chemotherapy. B, performance of different comorbidity measures where outcome is noncancer death within 24 months of diagnosis.

TABLE 1

Characteristics of the Study Population

	Overall Cohort (n = 5777)		Noncancer Death Cohort (n = 4003)	
	Number	Percentage*	Number	Percentage*
Age				
66–70	1247	21.6	916	22.9
71–75	1506	26.1	1046	26.1
76–80	1403	24.3	966	24.1
81–85	994	17.2	664	16.6
86+	627	10.9	411	10.3
Race				
Black	423	7.3	282	7.0
Asian	80	1.4	55	1.4
Hispanic	120	2.1	87	2.2
Other	283	4.9	200	5.0
White	4871	84.3	3379	84.4
Gender				
Male	2577	44.6	1785	44.6
Female	3200	55.4	2218	55.4
Marital status				
Married	3074	53.2	2177	54.4
Separated/divorced	663	11.5	452	11.3
Widowed	1941	33.6	1303	32.6
Missing	99	1.7	71	1.8
Race/age-specific median income in census tract				
<\$20,000	1630	28.2	1071	26.8
\$20,001–25,000	1068	18.5	766	19.1
\$25,001–30,000	898	15.5	642	16.0
>\$30,000	1675	29.0	1182	29.5
Missing	506	8.8	342	8.5
Race-specific percentage of high school graduates in census tract				
0–50%	123	2.1	80	2.0
51–75%	1613	27.9	1077	26.9
76–100%	4041	70.0	2846	71.1
Registry				
Connecticut	897	15.5	636	15.9
Georgia	331	5.7	225	5.6
Hawaii	145	2.5	115	2.9

	Overall Cohort (n = 5777)		Noncancer Death Cohort (n = 4003)	
	Number	Percentage*	Number	Percentage*
Iowa	935	16.2	657	16.4
Los Angeles	846	14.6	600	15.0
Michigan	944	16.3	641	16.0
New Mexico	185	3.2	122	3.0
San Francisco	501	8.7	327	8.2
San Jose	249	4.3	179	4.5
Seattle	522	9.0	355	8.9
Utah	222	3.8	146	3.6
Vital status				
Alive	3464	60.0	3464	86.5
Dead, cause not cancer	539	9.3	539	13.5
Dead, cause cancer	1704	29.5	0	0
Dead, cause unknown	70	1.2	0	0

* Percentages may not sum to 100% because of rounding errors.

TABLE 2
Frequencies of Different Comorbidity Measures: Crude Rates and Adjusted Odds Ratios (ORs) of Study Outcomes

Comorbidity Variables	Receipt of Chemotherapy*				Noncancer Death†					
	N (n = 5777)	%	Crude Rate	Adjusted OR (n = 5181)	95% CI	N (n = 4003)	%	Crude Rate	Adjusted OR (n = 3596)	95% CI
Elixhauser‡										
PrIP congestive heart failure	166	2.9	0.271	0.62	(0.4–0.98)	120	3.0	0.383	3.05	(1.8–5.16)
PrIP paralysis	40	0.7	0.325	0.70	(0.29–1.69)	25	0.3	0.440	5.29	(1.74–16.12)
PrIP neurological disorder	47	0.8	0.298	0.44	(0.19–0.99)	32	0.8	0.313	1.53	(0.56–4.21)
PrIP coagulopathy	7	0.1	0.143	0.28	(0.03–2.94)	5	0.1	0.400	16.02	(1.65–155.26)
PrIP psychosis	31	0.5	0.226	0.35	(0.13–0.94)	22	0.6	0.273	1.51	(0.46–4.93)
PrOP congestive heart failure	170	2.9	0.329	0.57	(0.39–0.85)	124	3.1	0.323	2.45	(1.54–3.91)
PrOP peripheral vascular disease	156	2.7	0.385	0.82	(0.55–1.22)	109	2.7	0.330	2.00	(1.23–3.26)
PrOP paralysis	13	0.2	0.231	0.12	(0.02–0.68)	12	0.3	0.333	3.56	(0.8–15.72)
PrOP neurological disorder	55	1.0	0.255	0.28	(0.13–0.58)	41	1.0	0.293	1.99	(0.84–4.71)
PrOP hypertension	1123	19.4	0.578	1.37	(1.15–1.63)	812	20.3	0.127	0.69	(0.52–0.92)
PrOP diabetes	393	6.8	0.552	0.85	(0.65–1.11)	284	7.1	0.197	1.75	(1.2–2.56)
ResIP hypertension	834	14.4	0.598	1.35	(1.12–1.64)	618	15.4	0.091	0.54	(0.39–0.77)
ResIP diabetes with chronic complications	64	1.1	0.484	0.79	(0.43–1.45)	47	1.2	0.404	3.93	(1.97–7.86)
ResIP psychosis	21	0.4	0.333	0.63	(0.19–2.04)	13	0.3	0.385	4.98	(1.3–19.12)
Klabunde score										
PrIP	5777	NA	NA	0.50	(0.40–0.63)	4003	NA	NA	2.60	(2.01–3.36)
PrOP	5777	NA	NA	0.95	(0.88–1.01)	4003	NA	NA	2.53	(1.85–3.45)
ResIP	5777	NA	NA	0.62	(0.45–0.87)	4003	NA	NA	1.91	(1.20–3.02)
DCG predicted expenditures										
\$1000 to \$1499	661	11.4	0.823	Ref		484	12.1	0.035	Ref	
\$1500 to \$1999	973	16.8	0.764	0.89	(0.64–1.24)	678	16.9	0.050	1.28	(0.63–2.59)
\$2000 to \$2499	615	10.7	0.662	0.92	(0.65–1.31)	443	11.1	0.056	1.13	(0.53–2.41)

Comorbidity Variables	Receipt of Chemotherapy [*]					Noncancer Death [†]				
	N (n = 5777)	%	Crude Rate	Adjusted OR (n = 5181)	95% CI	N (n = 4003)	%	Crude Rate	Adjusted OR (n = 3596)	95% CI
\$2500 to \$2999	985	17.1	0.571	0.94	(0.67–1.33)	657	16.4	0.129	2.39	(1.23–4.67)
\$3000 to \$3999	1137	19.7	0.391	0.58	(0.42–0.82)	784	19.6	0.175	3.09	(1.61–5.96)
\$4000 to \$4999	476	8.2	0.429	0.57	(0.39–0.82)	338	8.4	0.178	3.30	(1.67–6.53)
\$5000 to \$5999	292	5.1	0.329	0.36	(0.24–0.55)	195	4.9	0.251	5.18	(2.55–10.49)
\$6000 to \$7499	261	4.5	0.349	0.44	(0.29–0.67)	168	4.2	0.274	5.87	(2.87–12.03)
\$7500 to \$9999	187	3.2	0.337	0.38	(0.24–0.61)	127	3.2	0.291	5.35	(2.54–11.29)
\$10,000+	190	3.3	0.300	0.31	(0.19–0.49)	129	3.2	0.380	9.02	(4.37–18.62)
Count of major ADGs										
0	3336	57.7	0.604	Ref		2310	57.7	0.087	Ref	
1	1541	26.7	0.533	0.81	(0.7–0.94)	1077	26.9	0.164	1.91	(1.51–2.41)
2	617	10.7	0.418	0.55	(0.45–0.69)	431	10.8	0.260	2.99	(2.25–3.97)
3+	283	4.9	0.410	0.55	(0.41–0.74)	185	4.6	0.260	2.84	(1.91–4.24)

^{*} N for crude rate of receipt of chemotherapy is 5777, N for regression is 5181 because 99 were excluded due to unknown marital status, and 497 were excluded due to unknown race-specific median income.

[†] N for crude rate or noncancer deaths is 4003, N for regression is 3596 because 71 were excluded due to unknown marital status and 336 were excluded due to unknown race-specific median income.

[‡] Only significant Elixhauser comorbidity indicators are included in Table 2. Reference category of each Elixhauser indicator is not having the condition.

PrIP indicates prior inpatient claims; PrOP, prior outpatient claims; ResIP, resection inpatient claims.