

Medicare & Medicaid Research Review
2012: Volume 2, Number 3

*A publication of the Centers for Medicare & Medicaid Services,
Center for Strategic Planning*

Evaluating Comorbidity Scores Based on Health Service Expenditures

Meredith L. Kilgore,¹ Wilson Smith,¹ Jeffrey R. Curtis,² Michael A. Morrissey,¹ David J. Becker,¹
Kenneth G. Saag,² Elizabeth Delzell¹

¹UAB—School of Public Health

²UAB—School of Medicine

Objective: To describe the performance of Charlson Comorbidity Index (CCI) specifications among Medicare beneficiaries and subgroups.

Data Sources: Medicare data for beneficiaries covered by Parts A and B and not Medicare Advantage throughout 2007.

Study Design: We evaluated several CCI specifications, particularly a model using expenditures related to Charlson categories, to predict 1 year mortality.

Data Collection/Extraction Methods: Data were obtained from the Chronic Condition Data Warehouse.

Principal Findings: The use of Charlson related expenditures did not result in improved mortality prediction. CCI models perform less well in population subgroups with higher underlying mortality risks based on age and chronic conditions.

Conclusions: Relatively simple models provide quite adequate discrimination compared to more sophisticated models. Our proposed and more sophisticated model, which added in expenditure information, did not perform as well as much more easily executed methods.

Keyword: Comorbidity scores; risk adjustment; Medicare claims data

doi: <http://dx.doi.org/10.5600/mmrr.002.03.a05>

Introduction

Much ink has been spilled on using comorbidity scores for risk adjustment, most often addressing methods for adapting the Charlson Comorbidity Index (CCI) to administrative data (Deyo, Cherkin, & Ciol, 1992; Romano, Roos, & Jollis, 1993; Schneeweiss, Wang, Avorn, & Glynn, 2003; Schneeweiss et al., 2004; Klabunde et al., 2007; Gagne et al., 2011). Some questions remain open as to how CCI methods compare in different populations, for different outcomes (e.g., mortality, hospitalization), the optimal weights, and on the utility and best method of including outpatient diagnoses in the index. In some studies, there has been lack of clarity as to precisely which types of claims should be included in estimating mortality risk.

This report compares the performance of a new comorbidity index, using health expenditures associated with the Charlson disease categories, to existing methods used for calculating the CCI. We conjectured that using a weighting scheme based on expenditures would outperform the conventional Charlson weights which are computed from a set of comorbidity-derived indicators. The basis of this conjecture was that expenditures would provide a measure of the relative severity of disease—individuals with very mild diabetes might be on medications, but may have only a few office visits related to the condition, while those with more severe disease would have many more visits, possibly including hospitalizations, which would result in higher levels of expenditure. Thus, we hypothesized that expenditures could provide a weighting scheme to differentiate individuals with varying levels of disease severity. We also extend the work published by Schneeweiss and colleagues (2003, 2004), by comparing various approaches to CCI calculation and the performance of the CCI in different subgroups of Medicare beneficiaries facing different underlying mortality risks.

Methods

After receiving approval for this study from the Institutional Review Board at the University of Alabama at Birmingham, we obtained Medicare data from the Chronic Condition Data Warehouse (Buccaneer Computer Systems & Service, Inc, West Des Moines, IA). We used enrollment and claims data for a 5% sample of Medicare beneficiaries, 65 years of age and older, who were covered under both Parts A and B throughout the year 2007; we excluded beneficiaries who were enrolled in a Medicare Advantage plan at any time during the year because their claims were incompletely recorded in the data system. The data were split randomly into two equal samples, a “training” sample used to calculate weights for deriving comorbidity scores and a validation sample used to assess the discrimination capacity for each of the prediction models being evaluated.

We used the approach described by Romano et al (1993), and updated by Quan, Parsons, & Ghali (2002) to identify CCI categories (myocardial infarction, heart failure, peripheral vascular

disease, cerebrovascular disease, dementia, chronic pulmonary disease, connective tissue disease, gastrointestinal ulcers, liver disease, kidney disease, diabetes mellitus, hemiplegia, malignancy, metastatic neoplasms, autoimmune deficiency syndrome). The precise coding algorithms and coding program are available on request. Indicators for each CCI category were derived separately from inpatient and outpatient (physician and outpatient hospital) claims and CCI scores were calculated.

We calculated CCI scores for beneficiaries who were alive at the beginning of 2008 based on claims filed throughout 2007. Our CCI condition indicators were constructed using primary and secondary diagnosis codes from inpatient and outpatient hospital claims, along with diagnosis codes from physician (carrier) claims containing physician encounter codes. To calculate CCI scores, we used the weights derived by Schneeweiss et al (2004) and also derived weights for a random selection comprising half of our Medicare beneficiaries. These weights were derived first by regressing an indicator for death at any time in 2008, on indicators' conditions identified in 2007 in the training sample, then assigned weights based on odds ratios for each indicator, rounded to the nearest integer. These weights were used to compute a CCI score for beneficiaries in the validation sample.

To construct our new index we also summed all payments for Medicare covered services (except for Part D) associated with the Charlson category diagnosis codes for each beneficiary. We used total payments (not charges) to reflect the best approximation of the costs of care. Thus the totals for each indicator category consisted of Medicare payments to providers, plus any coinsurance or deductible payments and any payments made by other payers when Medicare was not the primary payer. Thus, the new CCI score is the sum total of expenditures across all Charlson categories. To evaluate the effect of expenditures on odds ratios for 1-year mortality, we measure expenditures in \$10,000 units.

The discrimination of each CCI measure was compared to a baseline model containing age, gender, and race/ethnicity with or without an indicator for any previous hospitalization. We provided separate analyses for all beneficiaries and for only those who were hospitalized during the baseline period, because both approaches are used in the literature. Some analyses use only inpatient claims or discharge data. We used c-statistics from logistic regressions to evaluate CCI performance (Hosmer et al., 1997). C-statistics range from 0.5, representing a complete lack of discrimination, to 1.0, representing perfect discrimination for dichotomous events (in this case mortality any time in 2008). In general, a c-Statistic of 0.7 or more represents adequate discrimination, 0.8 is very good, and 0.9 or more is excellent (and seldom seen, Liebetrau, 1983).

Finally, we stratified the sample by age group, gender, race, and by Chronic Condition Data Warehouse defined conditions (Buccaneer, Inc., 2010). We then compared discrimination performance using the best overall model for each of the subgroups. All data management and statistical analyses were conducted using SAS for Windows Version 9.1.3 (SAS Institute, Inc., Cary, NC).

Results

Exhibit 1 presents c-statistics for various model specifications, first among beneficiaries with at least one inpatient stay in 2007, then for all beneficiaries. The first panel is presented, because a number of analyses are confined to inpatient data only. Among this group, the model containing individual CCI indicators slightly outperforms the model containing only the composite score ($c=0.741$, 95% CI 0.739–0.743, versus $c=0.731$, 95%

Exhibit 1. c-Statistics for Comorbidity Models

Model Specification	C-Statistic	(95% CI)
Beneficiaries with an Inpatient Stay (N = 189,654 Deaths = 28,185)		
Age, Race, Sex	0.643	(0.641–0.645)
+ Charlson Indicators	0.741	(0.739–0.743)
+ Charlson Score	0.731	(0.729–0.733)
All Beneficiaries (N = 1,083,781 Deaths = 64,219)		
Age, Race, Sex	0.715	(0.714–0.716)
+ Indicator for Any Inpatient Stay	0.767	(0.766–0.768)
+ Charlson Indicators (inpatient + outpatient claims)	0.804	(0.803–0.805)
+ Charlson Score (inpatient + outpatient claims)	0.800	(0.799–0.801)
+ Charlson Indicators (inpatient claims only)	0.746	(0.745–0.747)
+ Charlson Score (inpatient claims only)	0.779	(0.778–0.780)
+ Charlson Indicators (inpatient + outpatient claims)	0.800	(0.799–0.801)
+ Charlson Score (inpatient + outpatient claims)	0.796	(0.795–0.797)
+ Charlson Expenditures	0.749	(0.748–0.750)

SOURCE: Medicare Claims and Enrollment Files, 2007–2008. All of our analyses use 2007 data to predict death in 2008.

CI 0.729–0.733), but if there are limited degrees of freedom, the score provides adequate discrimination.

For all beneficiaries, simply including an indicator for any hospital stay dramatically improves the discrimination achieved with age, race, and gender ($c=0.767$, 95% CI 0.766–0.768), and adding CCI indicators or scores adds further discrimination ($c=0.804$, 95% CI 0.803–0.805 and $c=0.800$, 95% CI 0.799–0.801, respectively). CCI indicators and scores based on inpatient claims alone provided lower discrimination ($c=0.746$, 95% CI 0.745–0.747 and $c=0.779$, 95% CI 0.778–0.780), but were still within a range considered adequate. Adding outpatient claims data in deriving CCI increased the discrimination to almost the same degree found when used in common with the inpatient stay indicator. The use of expenditures to capture disease severity failed to outperform the other methods ($c=0.749$, 95% CI 0.748–0.750). Each \$10,000 in expenditures was associated with an increase of 1.80 in the OR for 1-year mortality ($p < 0.0001$).

We found no difference in performance between the weights derived from our split sample and the weights published previously (Schneeweiss, et al., 2004), so the results from those models are not presented separately. (Our weights were 7 for metastatic cancer, 5 for dementia, 3 for congestive heart failure, 2 for chronic obstructive pulmonary disease, hemiplegia, chronic kidney disease, cancer, and liver disease and 1 for all other comorbid conditions.)

Exhibit 2. Model Performance Variation in Beneficiary Subgroups

Stratification	c-Statistic	95% CI	CCW* Condition	c-Statistic	95% CI
Age Group			Myocardial Infarction	0.754	0.745–0.763
65<=age<=74	0.774	0.773–0.775	Alzheimer's Disease	0.666	0.662–0.670
75<=age<=84	0.742	0.741–0.743	Atrial Fibrillation	0.761	0.758–0.764
85<=age	0.683	0.680–0.686	Cataracts	0.809	0.808–0.810
Gender			Chronic Kidney Disease	0.745	0.743–0.747
Women	0.810	0.809–0.811	COPD	0.730	0.727–0.733
Men	0.787	0.786–0.788	Diabetes	0.788	0.786–0.790
Race/Ethnicity			Glaucoma	0.810	0.808–0.812
White	0.802	0.801–0.803	Hip Fracture	0.723	0.714–0.732
Black	0.787	0.784–0.790	Ischemic Heart Disease	0.783	0.782–0.784
Hispanic	0.793	0.789–0.797	Osteoporosis	0.818	0.816–0.820
Other	0.804	0.800–0.808	Rheumatoid/Osteoarthritis	0.808	0.806–0.810
			Stroke/TIA	0.742	0.738–0.746
			Breast Cancer	0.800	0.795–0.805
			Colorectal Cancer	0.764	0.756–0.772
			Prostate Cancer	0.800	0.796–0.804
			Lung Cancer	0.716	0.706–0.726
			–	0.793	0.774–0.812

*Chronic Conditions Warehouse

SOURCE: Medicare Claims and Enrollment Files, 2007–2008. All of our analyses use 2007 data to predict death in 2008.

Exhibit 2 shows the performance of the best model (demographics, hospital stay indicator, and CCI indicators derived from inpatient and outpatient claims) in different subgroups of Medicare beneficiaries. Discrimination was lower for the oldest old, for men compared with women, for Blacks and Hispanics compared with Whites and Others, and for individuals with more serious illnesses. The lowest level of discrimination was for individuals with Alzheimer's Disease ($c=0.666$), the only instance of a c-statistic below 0.70 (generally consider the threshold for useful discrimination).

Conclusions

Our new, more nuanced (and difficult to execute), variant of the Charlson score failed to perform as hoped. We tested a log-transformation of expenditures to no better result. In general, most approaches were reasonable predictors of mortality, and simple models were often quite good. A simple indicator of any inpatient stay had better discrimination than CCI indicators

derived from inpatient claims only. When models are based solely on inpatient claims, the Charlson score outperforms the indicators. Adding outpatient claims to the derivation did achieve better discrimination, but did not provide a great deal of return for the extra effort. Our findings suggest that more information on patient comorbidity may not provide much return in predicting one year survival for risk adjustment, implying that there are many other factors that affect survival rates than can readily be captured *a priori*.

Our subgroup analysis indicates that the discriminatory power of this approach varies depending on the population in which it is applied. Specifically, discrimination is inversely proportional to the baseline hazard. The model performed less well among the very old, at risk minorities, and the very ill.

Correspondence

Meredith L. Kilgore, Ph.D., UAB—School of Public Health, 1530 3rd Ave. South, Birmingham, AL 35294, mkilgore@uab.edu, Tel. (205) 975-8840, Fax. 205-934-3347.

Financial Disclosure

This research was supported by a contract between UAB and Amgen, Inc. Only the authors from UAB had access to the Medicare data used.

References

- Buccaneer, Inc. (2010, October), CMS Chronic Condition Data Warehouse Condition Categories. Retrieved from http://www.ccwdata.org/cs/groups/public/documents/document/ccw_conditioncategories.pdf, accessed 06/07/2011.
- Deyo, R. A., Cherkin, D. C., & Ciol, M. A. (1992, June). Adapting a clinical comorbidity index for use with ICD9-CM administrative databases. *Journal of Clinical Epidemiology*, *45*(6), 613–619. [PubMed http://dx.doi.org/10.1016/0895-4356\(92\)90133-8](http://dx.doi.org/10.1016/0895-4356(92)90133-8)
- Gagne, J. J., Glynn, R. J., Avorn, J., Levin, R., & Schneeweiss, S. (2011, July). A combined comorbidity score predicted mortality in elderly patients better than existing scores. *Journal of Clinical Epidemiology*, *64*, 749–759. [PubMed http://dx.doi.org/10.1016/j.jclinepi.2010.10.004](http://dx.doi.org/10.1016/j.jclinepi.2010.10.004)
- Hosmer, D. W., Hosmer, T., Le Cessie, S., & Lemeshow, S. (1997, May). A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*, *16*(9), 965–980. [PubMed http://dx.doi.org/10.1002/\(SICI\)1097-0258\(19970515\)16:9<965::AID-SIM509>3.0.CO;2-O](http://dx.doi.org/10.1002/(SICI)1097-0258(19970515)16:9<965::AID-SIM509>3.0.CO;2-O)
- Klabunde, C. N., Legler, J. M., Warren, J. L., Baldwin, L. M., & Schrag, D. (2007, August). A refined comorbidity measurement algorithm for claims-based studies of breast, prostate, colorectal, and lung cancer patients. *Annals of Epidemiology*, *17*(8), 584–590. [PubMed http://dx.doi.org/10.1016/j.annepidem.2007.03.011](http://dx.doi.org/10.1016/j.annepidem.2007.03.011)
- Liebetrau, A. M. *Measures of Association. Quantitative Application in the Social Sciences*. Vol. 32. Beverly Hills, Calif: Sage Publications; 1983.
- Quan, H., Parsons, G. A., & Ghali, W. A. (2002, August). Validity of Information on Comorbidity Derived From ICD-9-CCM Administrative Data. *Medical Care*, *40*, 675–685. [PubMed http://dx.doi.org/10.1097/00005650-200208000-00007](http://dx.doi.org/10.1097/00005650-200208000-00007)
- Romano, P. S., Roos, L. L., & Jollis, J. G. (1993, October). Adapting a clinical comorbidity index for use with ICD9-CM administrative data: Differing perspectives. *Journal of Clinical Epidemiology*, *46*, 1075–1079. [PubMed http://dx.doi.org/10.1016/0895-4356\(93\)90103-8](http://dx.doi.org/10.1016/0895-4356(93)90103-8)
- Schneeweiss, S., Wang, P. S., Avorn, J., & Glynn, R. J. (2003, August). Improved comorbidity adjustment for predicting mortality in Medicare populations. *Health Services Research*, *38*(4), 1103–1120. [PubMed http://dx.doi.org/10.1111/1475-6773.00165](http://dx.doi.org/10.1111/1475-6773.00165)
- Schneeweiss, S., Wang, P. S., Avorn, J., Maclure, M., Levin, R., & Glynn, R. J. (2004, May). Consistency of performance ranking of comorbidity adjustment scores in Canadian and U.S. utilization data. *Journal of General Internal Medicine*, *19*, 444–450. [PubMed http://dx.doi.org/10.1111/j.1525-1497.2004.30109.x](http://dx.doi.org/10.1111/j.1525-1497.2004.30109.x)

Medicare & Medicaid Research Review
2012
Volume 2, Number 3

Mission Statement

Medicare & Medicaid Research Review is a peer-reviewed, online journal reporting data and research that informs current and future directions of the Medicare, Medicaid, and Children's Health Insurance programs. The journal seeks to examine and evaluate health care coverage, quality and access to care for beneficiaries, and payment for health services.

<http://www.cms.gov/MMRR/>

U.S. Department of Health & Human Services

Kathleen Sebelius
Secretary

Centers for Medicare & Medicaid Services

Marilyn Tavenner
Acting Administrator

Editor-in-Chief David M. Bott, Ph.D.

Senior Editor Cynthia Riegler, M.A.

Associate Editors

John Hsu, M.D., M.B.A, M.S.C.E.
Harvard Medical School
James H. Marton, Ph.D
Georgia State University

Jennifer Polinski, Sc.D, M.P.H.
Brigham & Women's Hospital
Robert Weech-Maldonado, Ph.D.
University of Alabama at Birmingham

Editorial Board

Gerald S. Adler, M.Phil.
CMS/Center for Strategic Planning
Andrew Bindman, M.D.
University of California, San Francisco
William J. Buczko, Ph.D.
CMS/Innovation Center
Todd Caldis, Ph.D., J.D.
CMS/Office of the Actuary
Craig F. Caplan, M.A.
CMS/ Center for Medicare

Melissa A. Evans, Ph.D.
CMS/Center for Program Integrity
Jesse M. Levy, Ph.D.
CMS/ Innovation Center
Isidor R. Strauss, F.S.A.
CMS/Office of the Actuary
Fred G. Thomas, Ph.D., C.P.A.
CMS/ Innovation Center
Jill Boylston Herndon, Ph.D.
University of Florida

Contact: mmrr-editors@cms.hhs.gov

Published by the Centers for Medicare & Medicaid Services

All material in the Medicare & Medicaid Research Review is in the public domain and may be duplicated without permission.

Citation to source is requested.