Published in final edited form as:

J Geriatr Oncol. 2013 April 1; 4(2): 157–165. doi:10.1016/j.jgo.2012.12.005.

A Novel Approach to Improve Health Status Measurement in Observational Claims-based Studies of Cancer Treatment and Outcomes

Amy J. Davidoff, M.S., Ph.D.^a, Arti Hurria, M.D.^d, Ilene H. Zuckerman, PharmD, Ph.D.^b, Stuart M. Lichtman, M.D.^e, Naimish Pandya, M.D.^c, Arif Hussain, M.D.^c, Franklin Hendrick, B.S.^b, Dr. Jonathan P Weiner, P.H.^f, Xuehua Ke, M.A.^b, and Martin J. Edelman, M.D.^c

^aAgency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, USA

^bUniversity of Maryland School of Pharmacy, 220 Arch Street, 12th floor, Baltimore, MD 2120, USA

^cUniversity of Maryland Greenebaum Cancer Center, 22 S. Greene Street Baltimore, MD 21201, USA

^dCity of Hope, 1500 E. Duarte Road, Duarte, CA 91010, USA

^eMemorial Sloan-Kettering Cancer Center, 650 Commack Road, Commack, New York 11725, USA

^fJohns Hopkins Bloomberg School of Public Health, 624 N. Broadway, Room 605, Baltimore, MD 21205, USA

Abstract

Objectives—To develop and provide initial validation for a multivariate, claims-based prediction model for disability status (DS), a proxy measure of performance status (PS), among older adults. The model was designed to augment information on health status at the point of cancer diagnosis in studies using insurance claims to examine cancer treatment and outcomes.

Materials and Methods—We used data from the 2001–2005 Medicare Current Beneficiary Survey (MCBS), with observations randomly split into estimation and validation subsamples. We developed an algorithm linking self-reported functional status measures to a DS scale, a proxy for the Eastern Cooperative Oncology Group (ECOG) PS scale. The DS measure was dichotomized to

© 2013 European Society for Clinical Nutrition and Metabolism. Published by Elsevier Ltd. All rights reserved.

Corresponding author information/address for reprints: Franklin Hendrick, Pharmaceutical Health Services Research, University of Maryland School of Pharmacy, 220 Arch Street, 12th Floor, Baltimore, MD 21201, fhend001@umaryland.edu, Telephone: 301-919-0417, Fax: 410-706-5394.

Amy J. Davidoff, M.S., Ph.D.^a, amy.davidoff@ahrq.hhs.gov, Phone: 301-427-1678, Fax: 301-427-

Arti Hurria, M.D.^d, ahurria@coh.org, Phone: 626-256-4673 ext. 64173, Fax: 626-301-8898

Ilene H. Zuckerman, PharmD, Ph.D.^b, izuckerm@rx.umaryland.edu, Phone: 410-706-3266, Fax: 410-706-5394

Stuart M. Lichtman, M.D.^e, lichtmas@mskcc.org, Phone: 631-623-4100, Fax: 631-864-3827

Naimish Pandya, M.D.^c, npandya@umm.edu, Phone: 410-328-2567, Fax: 410-328-6896

Arif Hussain, M.D.^c, ahussain@som.umaryland.edu, Phone: 410-328-7225, Fax: 410-328-0805

Franklin Hendrick, B.S.^b, fhend001@umaryland.edu, Phone: 301-919-0417, Fax: 410-706-5394

Jonathan P. Weiner, Dr. P.H.^f, jweiner@jhsph.edu, Phone: 410-955-5661, Fax: 410-955-0470

Xuehua Ke, M.A.^b, xke001@umaryland.edu, Phone: 410-706-1418, Fax: 410-706-5394

Martin J. Edelman, M.D.^c, medelman@umm.edu, Phone: 410-328-2703, Fax: 410-328-0805

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

focus on good [ECOG 0–2] versus poor [ECOG 3–4] PS. We identified potential claims-based predictors, and estimated multivariate logistic regression models, with poor DS as the dependent measure, using a stepwise approach to select the optimal model. Construct validity was tested by determining whether the predicted DS measure generated by the model was a significant predictor of survival within a validation sample from the MCBS.

Results and Conclusion—One-tenth of beneficiaries met the definition for poor DS. The base model yielded high sensitivity (0.79) and specificity (0.92); positive predictive value=48.3% and negative predictive value=97.8%, c-statistic=0.92 and good model calibration. Adjusted poor claims-based DS was associated with an increased hazard of death (HR=3.53, 95% CI 3.18, 3.92). The ability to assess DS should improve covariate control and reduce indication bias in observational studies of cancer treatment and outcomes based on insurance claims.

Keywords

Medicare; cancer; health status; observational study; treatment selection bias; functional status

Introduction

Observational studies using administrative data are increasingly used to provide information on population-based patterns of cancer treatment and to evaluate treatment outcomes through comparative effectiveness research.^{1, 2} For older adults and non-elderly adults with disabilities, the Surveillance, Epidemiology and End Results (SEER) registry, linked to Medicare enrollment and claims data, has become an important resource for this research, as have other claims based sources.^{3, 4} One challenge for researchers using administrative data is that treatment decisions integrate a variety of factors, including patient health status and patient or physician preferences and attitudes, not all of which can be measured fully in the available data. Most claims-based measures of health status rely on diagnosis codes to control for the presence of comorbidities at the time of cancer diagnosis.⁵ For example, diagnostic information has been used to create weighted indices, as in the case of the Charlson Comorbidity Index (CCI).⁶ However, dimensions of health status such as functional or performance status (PS), are difficult to characterize using the available methods.^{7, 8} To the extent that these poorly measured or unobserved factors are important determinants of treatment and survival, failure to take them into account can result in biased estimates.⁹

PS is a measure of patient functional capacity, with an emphasis on physical dimensions. PS incorporates the ability to work, time out of bed, and the ability to perform “self care.” An initial scale developed by Karnofsky was subsequently modified by the Eastern Cooperative Oncology Group (ECOG), and eventually adopted by the World Health Organization.^{10, 11} The scale is summarized in Appendix A. PS scores are commonly used throughout oncology practice as a general numerical guide to the cancer patient’s health. PS is assessed based on patient or proxy report of activity levels, combined with clinician observation of the patient’s mobility during a medical encounter. As a result of its prognostic value for survival, PS is used as a criterion for selection into clinical trials and as a key factor determining whether to actively treat cancer patients or provide supportive care. Treatment guidelines, for example, those promulgated by expert panels through the National Comprehensive Cancer Network, routinely tailor recommendations by patient PS.¹² PS is usually assessed at diagnosis to determine an initial treatment strategy, and may be updated regularly as treatment response and disease progression require reassessment of treatment.¹³ Furthermore, most cancer clinical trials are restricted to patients with ECOG PS score 0–2, with many further restricted to patients with PS score 0–1 because patients with poor PS

(≥ 3) are more likely to experience unacceptable toxicities and/or are less likely to experience survival benefit.¹⁴

Given the clinical relevance of PS in cancer patients, and the current limited ability to assess PS with existing claims-based comorbidity measures, we undertook the development and validation of a multivariate prediction model based on administrative claims to capture this dimension of health status. The resulting model can be used to augment health status information in research using administrative databases that lack PS or functional status information, but include administrative claims needed to operationalize the independent measures in the model. As we did not have direct measures of PS in the data used to develop the model, we first created a proxy measure which we refer to as disability status (DS) based on combinations of self-reported functional status measures.¹⁵ We describe the process of DS construction, model development, and initial validation steps for the DS model.

Materials and methods

We used data from the Medicare Current Beneficiary Survey (MCBS), a nationally representative rotating panel survey of community-based and institutionalized Medicare beneficiaries.¹⁶ The MCBS samples approximately 5,100 new beneficiaries each year, with up to four years of observation. The survey captures information on demographics, insurance, and self-reported health and functional status, including limitations and dependence in activities of daily living (ADLs) and instrumental activities of daily living (IADLs), difficulty with activities requiring strength, stamina or agility, and participation in exercise. The MCBS is linked by the Centers for Medicare and Medicaid Services (CMS) to Medicare Parts A and B claims for participants in each survey year, which provide detailed information concerning service date, provider, type of service, and associated diagnoses. Specific procedures are reported using either ICD-9-CM procedure codes, the American Medical Association's Current Procedural Terminology (CPT) codes, or the CMS Healthcare Common Procedure Coding System (HCPCS level II) codes. We used data from 2001, 2003, and 2005, as these years included biannual questions on exercise. We selected one observation-year per beneficiary. Beneficiaries with any Medicare Advantage enrollment or who resided in facilities other than assisted living or nursing homes (for example, group homes for persons with mental disabilities) were also excluded, the latter due to limited functional status information.

The initial sample included both older (age ≥ 65 years) and non-elderly adults with disabilities enrolled in Medicare. In sensitivity analyses we determined that model fit was substantially improved when we limited the sample to older adults. The final sample of Medicare beneficiaries was randomly split into estimation and validation samples, with 7,394 enrollees in each.

The MCBS is a sample of the general Medicare beneficiary population, not limited to patients with cancer. As a result, our model captures relationships between healthcare utilization and functional status, independent of the effects of cancer and its treatment. This is consistent with the approach used in the ECOG PS scale, and appropriate to capture baseline health status at the point of cancer diagnosis.

Key measures

Disability status—We developed a summary measure of DS to proxy for PS, based on self-reported measures of functional status, strength, stamina, and exercise, linked to the various functional dimensions and degrees of limitation specified in the ECOG PS scale. The approach, described in detail in Appendix B, was guided by a clinician panel representing a cross section of medical oncologists, including those with expertise in

geriatrics, and thoracic, gastrointestinal, breast, and prostate cancers. For each of the individual functional status measures we specified a common 5 point limitation scale (A='no difficulty', B='difficulty, no assistance needed/little difficulty/difficulty without help', C='difficulty, personal assistance available (on standby)/some difficulty', D='difficulty requiring equipment or assistance/a lot of difficulty', and E='can't do'). In our DS measure, we approximated the stepwise progression in the ECOG PS scale by grouping the functional status measures into those associated with heavy work, light work, and self-care activities, the dimensions used in the ECOG scale, and set thresholds for the number and degree of limitation associated with each level of disability. For example, the ECOG definition identifies patients with PS=3 as "capable of only limited self-care, confined to bed or chair more than 50% of waking hours," and patients with PS=4 as "completely disabled, unable to carry out any self-care, and totally confined to bed or chair." The DS algorithm required individuals with DS=3 to have limitations of at least level D in heavy work, in 3 of 5 measures of light work, and in dressing. Two out of three other measures of self-care were required to be at level C or greater. Requirements for DS=4 were substantially more restrictive. We created an algorithm that summarized information across the individual measures in each dimension, and then linked to the levels on the DS scale, assigning a survey-based DS level to each individual. We found that relatively few patients met the highest levels of activity limitations, so shifted thresholds to achieve greater variation in DS assignment. For the models reported in this paper, we collapsed the 5 level DS measure into a dichotomous indicator for good DS (DS=0–2) versus poor DS (DS= 3–4). The survey-based DS measures were used as the dependent variable for developing a prediction model using information from administrative claims as predictors.

Claims-based predictors—We identified a broad group of potential explanatory variables that included indicators for health care services, which were expected to vary based on DS level. These indicators were culled from three sources: services identified by clinicians in response to a series of patient vignettes ("What type of healthcare services would a patient with this degree of functional limitation be expected to use more or less commonly?"); variants of the Berenson-Eggers Type of Service (BETOS) codes; and selected demographic measures, which we tested for model stratification or interaction. The BETOS groups are comprised of CPT and HCPCS codes, and are intended to represent readily understood clinical categories, such as 'hip or knee replacement', which are stable over time.¹⁷ The BETOS indicators were pulled from physician and hospice claims during the calendar year in which the respondent's survey took place. We lumped or split selected BETOS categories to better capture the potential clinical relationship between the service and DS. Healthcare services specifically provided to patients with physical limitations, such as home oxygen, mobility aids (walker, wheelchair), or nursing home admission were expected to be associated with poor DS, while use of preventive services and elective surgical procedures were expected to be associated with better DS. Indicators were grouped into preventive services, evaluation and management (E&M) visits, other visit types, minor or ambulatory procedures, major procedures, imaging, durable medical equipment use, and other. We did not include indicators for chronic conditions or age, as DS is intended to capture a health status dimension independent of these factors.

Statistical analysis

We analyzed the distributions of the claims-based indicators overall and by survey-based DS (good/poor). Stepwise logistic regression predicting poor DS was used to select explanatory variables, using a 95% significance level for both variable entry and exit. We selected the optimal model as the one with the lowest Akaike information criterion (AIC), suggesting greatest model efficiency.^{18–21} Analyses were performed using SAS version 9.2 (SAS Institute, Inc. Cary, N.C.).

We estimated models with main effects only, and then permitted interaction terms to enter the stepwise regression modeling process, forcing interaction effects to remain in the model only if accompanied by the main effects. We permitted interactions with the health care service indicators and both region of the country and enrollment in Medicaid or Medicare Savings Programs ((MSP), where state programs pay Medicare Part B premiums), to capture the effects of differences in practice patterns that might be affected by region or supplemental insurance. Although indicators for all supplemental insurance types were available in the MCBS, only information on Medicaid/MSP participation is available in SEER-Medicare. In sensitivity analyses we used a three level measure of survey-based DS (0–1, 2, and 3–4) and estimated ordered logit models.

Model fit was tested by using the regression results to generate a predicted (claims-based) DS measure, converting the continuous prediction to a discrete indicator for (good/poor) and examining concordance-discordance between the predicted and survey-based DS.²² We examined the sensitivity and specificity associated with different cutpoints in the estimation sample, and selected the highest possible cutpoint that maintained a sensitivity of 80%. We chose this cutpoint as a reasonable approach to balance the trade-off between the false positive and negative rates of estimated poor DS. Model calibration was assessed with the Hosmer-Lemeshow goodness-of-fit test.²³ Model fit was tested both within the estimation and validation samples, and among the subgroup with poor self reported health status. Although the MCBS is designed to be representative of the Medicare population, there is oversampling of the oldest-old and younger beneficiaries with disabilities. We present unweighted analyses to be consistent with application in datasets where weights are not commonly employed; in sensitivity analyses we re-estimated the prediction models using MCBS sampling weights.

To further validate our DS measure we examined the relationship of both the survey-based and predicted claims-based DS measure with overall survival within the validation set of the MCBS. We calculated a follow-up period from October 1st (approximate timeframe for the Fall survey) of the relevant survey year until death. MCBS respondents, who did not die, were censored at December 31st, 2007. We estimated Cox proportional hazards (PH) models with either the survey-based or predicted DS, controlling for race, age, gender, educational attainment, marital status, household income, and residence location. We also included the CCI, to test whether the predicted DS measure had an independent association with survival. The project was approved by the University of Maryland Institutional Review Board.

Results

The characteristics of the estimation and prediction samples are provided in Table 1. Just under one tenth (9.3%) of beneficiaries met the definition for poor survey-based DS. Table 2 reports the prevalence by survey-based DS category of the claim indicator variables that appeared in either of the final models. The most common category of procedures was immunizations/vaccinations, reported for 45.8% of beneficiaries overall, with rates of 48.3% in the good DS group, and 21.5% in the poor DS group. Wheelchair claims were present for 4.9% overall, with rates of 3.4% of those with good DS and 19.5% among those with poor DS.

The estimated coefficients from the logistic regression models, with and without interactions between the service indicators and either Medicaid/MSP enrollment or region of the country, are provided in Appendix C. We do find that many relationships are consistent with clinical intuition, for example, that nursing home stays, home care, and ambulance use are associated with an increased probability of poor DS, while vaccinations, screenings, and cardiac monitoring and stress tests are associated with lower probability of poor DS. For the

model without interactions the c-statistic was .92, and the Hosmer-Lemeshow statistic indicated good model calibration. We identified cutpoints of 0.115 and 0.110 for models with and without interactions, respectively to generate the dichotomous indicator of good/poor predicted DS. Table 3 summarizes the within-sample concordance and out-of-sample predictive ability. Generally, the two models performed similarly. In the model without interactions within the validation half of the sample, the positive predictive value was 48.3% and the negative predictive value was 97.8%. Predictive value is a function of the underlying distribution of the dependent variable, survey-based DS. When we subset to those with self reported poor health status, the positive predictive value of predicted poor DS was improved (69.3%), while the negative predictive value was somewhat lower (84.2%). It is important to note that although the choice of probability cutpoints was chosen to maximize sensitivity and specificity, different situations might require focus on other aspects of model performance such as positive predictive value, and the cutpoints can be adjusted accordingly.

In sensitivity analyses using a three-category dependent variable (DS 0–1/2/ 3–4), we found that our model was unable to distinguish the middle group (DS=2) from the DS 0-1 group, resulting in poor specificity (results not shown). Weighted sensitivity analyses resulted in only negligible changes to the models and predictions (results not shown).

Table 4 provides the results of the Cox PH models examining the effect of the survey-based and predicted DS measures on overall survival. Median survival time for individuals with poor survey-based DS was 23 (95% CI 21, 27) months. This compared to 33 (95% CI 30, 36) months for individuals with poor claims-based DS. Median survival time was not reached in either the good survey-based DS or good claims-based DS individuals. In the adjusted models 2 and 3 we find the hazard ratio of death for individuals with poor survey-based DS (HR=3.97 (95% CI 3.55, 4.44)) is close to the hazard ratio of poor claims-based DS (HR=3.53 (95% CI 3.18, 3.92)). When the CCI is included in model 4 we find the claims-based DS point estimate is smaller but remains significant HR=2.85 [95% CI 2.56, 3.17].

Discussion

Our results demonstrate that healthcare service use indicators from administrative claims can be used to predict DS, and that the resulting predicted value is associated with survival in an older adult Medicare population. The explanatory variables were identified through a combination of clinician judgment as well as observed patterns within the data. As a result, the set of candidate variables were both clinically relevant and prevalent in the data. The explanatory variables in the final models reflect those associated with both good and poor DS. In a recent study examining the association between physician reported PS and claims-based service use measures in lung cancer patients (Salloum et al. 2011) and in selected studies examining determinants of cancer treatment and outcomes, indicators expected to be associated with poor PS were found to be associated with lower likelihood of aggressive cancer therapy.^{24–28} Inclusion of selected services that predict good DS is unique in this study. These latter measures were found to be significant in both bivariate comparisons and the stepwise regression modeling process.

Overall, we identified approximately 9% of older adult Medicare beneficiaries as having poor survey-based DS. There is little information available concerning PS of a broad population of cancer patients at diagnosis, since PS is not collected routinely unless patients are entered into clinical trials.²⁹ One study that examined ECOG PS among a consecutive series of 363 cancer patients reported that 67% of males and 80% of females present with good ECOG PS (PS of 0–1), with rates of good ECOG PS declining with increasing age.³⁰

We identified good survey-based DS in 90% of our overall sample, with 93% for males and 89% for females. Hence, our rates of good DS are somewhat higher, likely due to our inclusion of DS=2 in the “good” category, and that we are examining a general sample of older adults.

Our prediction models performed well in both the estimation and validation samples. However, the models did not perform as well among the Medicare enrolled non-elderly with disabilities. In exploratory analyses we found different distributions of claims indicators, especially lower use of preventive services among the adults with disabilities when compared to older adults reporting similar levels of functional status impairments. Conceivably, adults with long term disabilities may have different perceptions of their physical limitations, which may alter the relationship between reported functional status and healthcare service use.

The strong negative effect of both poor survey-based and predicted DS on overall survival provides initial validation that both measures capture the construct of PS. The slightly smaller magnitude of the claims-based predicted measure is to be expected – as there will be some error in the model-based prediction relative to the survey-based measure. Finally the retention of strong statistical significance after inclusion of the CCI in the survival models provides initial support for the independent significance of the DS measure, and the likely value in improving covariate control in cancer specific analyses of treatment and outcomes.

There are several limitations to our study that should be noted. Our dataset does not include a direct measure of ECOG PS. Instead, we constructed DS which serves as a proxy measure. Ideally we would have worked with a dataset that included both ECOG PS and the insurance claims. However, ECOG PS is routinely documented primarily in clinical trial records, which are usually specific to a single type of cancer, generally limited to patients with good ECOG PS, and not generally linked to claims. Hence, the data would not support the type of modeling question we address. The approach we used, relying on expert opinion and an iterative process to develop our DS proxy measure, should have resulted in a reasonable approximation for ECOG PS.

Our report of a predictive model to capture DS adds a new dimension to the use of diagnostic or procedure codes from insurance claims to capture and control for the effects of health status on treatment and outcomes. We have identified several areas for further validation of our model, and to explore how it may be used in research to examine comparative effectiveness, disparities, and quality. To further validate our model, we have implemented the DS prediction in previously developed SEER-Medicare analytic files and will test the ability to explain observed treatments and outcomes. We envision that this prediction model could also become a tool used in quality measurement. For example, it might be a mechanism to identify population level patterns of aggressive treatment for patients with poor DS, when there is a less favorable risk:benefit ratio. More general predictions of functional status may be useful in all of these roles, when applied to patients with chronic conditions other than cancer.

Cancer is the second leading cause of mortality in the U.S.³¹ The burden of cancer is particularly heavy for the elderly population, where cancer incidence rates are almost 10 times that of the under-65 population.³² Continued research to evaluate the benefits of cancer treatment in older patients is essential to public policymakers and medical insurers.³³ The ability to assess DS from claims data should significantly improve the validity of results from observational studies using administrative claims data such as SEER-Medicare. This research will help us identify how the risk/benefit ratio of therapy varies among patients with different DS status. This tool will also permit improved covariate control, and reduce

treatment selection bias in observational studies designed to examine the effectiveness of new therapies as they are adopted for treatment of older patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Dreyer NA, Tunis SR, Berger M, Ollendorf D, Mattox P, Gliklich R. Why observational studies should be among the tools used in comparative effectiveness research. *Health Aff (Millwood)*. 2010; 29:1818–1825. [PubMed: 20921481]
2. Fleurence RL, Naci H, Jansen JP. The critical role of observational evidence in comparative effectiveness research. *Health Aff (Millwood)*. 2010; 29:1826–1833. [PubMed: 20921482]
3. Potosky ALM, Riley GFM, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer related health services research using a linked Medicare-tumor registry database. *Med Care*. 1993; 31:749–756. [PubMed: 8336513]
4. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care*. 2002; 40(IV):3–18.
5. Klabunde CN, Warren JL, Legler JM. Assessing comorbidity using claims data: an overview. *Med Care*. 2002; 40(IV):26–35. [PubMed: 11748424]
6. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987; 40:373–383. [PubMed: 3558716]
7. Edwards BK, Brown ML, Wingo PA, Howe HL, Ward E, Ries LA, et al. Annual report to the nation on the status of cancer, 1975–2002, featuring population-based trends in cancer treatment. *J Natl Cancer Inst*. 2005; 97:1407–1427. [PubMed: 16204691]
8. Baldwin LM, Klabunde CN, Green P, Barlow W, Wright G. In search of the perfect comorbidity measure for use with administrative claims data: does it exist? *Med Care*. 2006; 44:745–753. [PubMed: 16862036]
9. Giordano SH, Kuo Y, Duan Z, Hortobagyi GN, Freeman J, Goodwin JS. Limits of observational data in determining outcomes from cancer therapy. *Cancer*. 2008; 112:2456–2466. [PubMed: 18428196]
10. Karnofsky, DA.; Burchenal, JH. The clinical evaluation of chemotherapeutic agents in cancer. In: MacLeod, CM., editor. *Evaluation of chemotherapeutic agents*. New York: Columbia University Press; 1949. p. 169
11. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982; 5:649–656. [PubMed: 7165009]
12. National Comprehensive Cancer Network. Clinical practice guidelines in oncology. 2008. <http://www.nccn.org>.
13. Deppermann KM. Influence of age and comorbidities on the chemotherapeutic management of lung cancer. *Lung Cancer*. 2001; 33:S115–S120. [PubMed: 11576716]
14. Lilenbaum RC, Herndon JE, List MA, Desch C, Watson DM, Miller AA, et al. Single-agent versus combination chemotherapy in advanced non-small-cell lung cancer: the Cancer and Leukemia Group B (study 9730). *J Clin Oncol*. 2005; 23:190–196. [PubMed: 15625373]
15. Salloum RG, Smith TJ, Jensen GA, Lafata JE. Using claims-based measures to predict performance status score in patients with lung cancer. *Cancer*. 2011; 117:1038–1048. [PubMed: 20957722]
16. Adler GS. A profile of the Medicare current beneficiary survey. *Health Care Financ R*. 1994; 15:153–163.
17. Centers for Medicare and Medicaid Services-Berenson-Eggers type of service (BETOS). 2010 https://www.cms.gov/HCPCSRReleaseCodeSets/20_BETOS.asp

18. Wang PS, Walker A, Tsuang M, Orav EJ, Levin R, Avorn J. Strategies for improving comorbidity measures based on Medicare and Medicaid claims data. *J Clin Epidemiol.* 2000; 53:571–578. [PubMed: 10880775]
19. Shtatland, ES.; Barton, MB.; Cain, E. SUGI '26 Proceeding, Paper 258-28. Cary, NC: SAS Institute, Inc; 2001. The perils of stepwise logistic regression and how to escape them using information criteria and the output delivery system.
20. Shtatland, ES.; Kleinman, K.; Cain, E. SUGI '29 Proceeding, Paper 191-29. Cary, NC: SAS Institute, Inc; 2002. A new strategy of model building in proc logistic with automatic variable selection, validation, shrinkage, and model averaging.
21. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J Roy Stat Soc B.* 2002; 64:583–639.
22. Lipsitz SR, Ritzmaurice GM, Molenberghs G. Goodness-of-fit tests for ordinal response regression models. *J Roy Stat Soc C-App.* 1996; 45:175–190.
23. Hosmer, DW.; Lemeshow, S. Applied logistic regression. 1st ed.. New York: John Wiley & Sons, Inc; 2000.
24. Davidoff AJ, Rapp T, Onukwugha E, Zuckerman IH, Hanna N, Pandya N, et al. Trends in disparities in receipt of adjuvant therapy for elderly stage III colon cancer patients: the role of the medical oncologist evaluation. *Med Care.* 2009; 47:1229–1236. [PubMed: 19786906]
25. Zuckerman IH, Rapp T, Onukwugha E, Davidoff A, Choti MA, Gardner J, et al. Effect of age on survival benefit of adjuvant chemotherapy in elderly patients with stage III colon cancer. *J Am Geriatr Soc.* 2009; 57:1403–1410. [PubMed: 19563521]
26. Davidoff AJ, Tang M, Seal B, Edelman MJ. Chemotherapy and survival benefit in elderly patients with advanced non-small cell lung cancer. *J Clin Oncol.* 2010; 28:2191–2197. [PubMed: 20351329]
27. Davidoff AJ, Gardner JF, Seal B, Edelman MJ. Population-based estimates of survival benefit associated with combined modality therapy in elderly patients with locally advanced non-small cell lung cancer. *J Thorac Oncol.* 2011; 6:934–941. [PubMed: 21372743]
28. Schneider M, Zuckerman IH, Onukwugha E, Pandya N, Seal B, Gardner J, et al. Chemotherapy treatment and survival in older women with estrogen receptor negative metastatic breast cancer: a population-based analysis. *J Am Geriatr Soc.* 2011; 59:637–646. [PubMed: 21453377]
29. Puts MTE, Hardt J, Monette J, Girre V, Springall E, Alibhai SMH. Use of geriatric assessment for older adults in the oncology setting: A systematic review. *J Natl Cancer Inst.* 2012; 104:1134–1164.
30. Repetto L, Fratino L, Audisio RA, Venturino A, Gianni W, Vercelli M, et al. Comprehensive geriatric assessment adds information to Eastern Cooperative Oncology Group performance status in elderly cancer patients: an Italian Group for Geriatric Oncology study. *J Clin Oncol.* 2002; 20:494–502. [PubMed: 11786579]
31. Minino AM, Heron MP, Murphy SL, Kochanek KD. Deaths: final data for 2004. *National Vital Statistics Reports* 2007. 2007; 55:1–119.
32. National Cancer Institute, DCCPS. Surveillance Research Program, Cancer Statistics Branch. Surveillance, Epidemiology, and End Results (SEER)*Stat database. 2010. <http://seer.cancer.gov/seerstat/>
33. Ramsey SD. How should we pay the piper when he's calling the tune? On the long-term affordability of cancer care in the United States. *J Clin Oncol.* 2007; 25:175–179. [PubMed: 17210936]

Table 1Characteristics of Medicare Beneficiaries in Estimation and Validation Samples, n=14,788^{a,b}

	Proportion of Sample	
	Estimation N=7,394	Validation N=7,394
Age		
65–69	16.4	16.5
70–74	22.2	22.8
75–79	21.0	20.4
80–84	20.3	20.9
85+	20.2	19.5
Race		
All others	2.5	2.8
White	88.0	88.1
Black	8.9	8.6
Hispanic	0.7	0.5
Sex		
Female	60.8	62.2
Male	39.2	37.8
Educational attainment		
Less than High School	33.4	34.5
High School Graduate	29.7	28.9
High School and Some College	22.1	22.1
College Graduate or Greater	14.9	14.6
Marital status		
Never Married	3.5	3.6
Married	48.3	48.2
Widowed	40.7	40.6
Divorced	6.7	6.8
Separated	0.9	0.8
Survey-based DS measure		
0,1,2	90.7	91.1
3,4	9.3	9.0
Residence location		
Rural	33.4	34.2
Urban	66.6	65.9
Year of data		
2001	35.5	35.1
2003	31.2	30.4
2005	33.3	34.6
Panel induction year		
1998	12.3	12.4

	Proportion of Sample	
	Estimation N=7,394	Validation N=7,394
1999	13.8	13.3
2000	17.2	17.0
2001	13.4	13.2
2002	16.5	16.5
2003	12.4	13.0
2004	14.5	14.6
Self-reported health status		
Excellent	13.73	14.19
Very Good	27.91	26.72
Good	31.8	33.6
Fair	19.0	18.5
Poor	7.3	6.6
Difficulty with ADL/IADL counts ^c		
0	45.3	46.0
1–3	26.4	25.6
4–5	10.8	11.1
6–7	5.6	5.2
8–10	11.9	12.1
Medicaid enrollment status		
No	84.9	84.4
Yes	15.1	15.6
Region of the country		
Northeast	16.8	17.8
Midwest	26.6	25.9
South	41.5	41.5
West	15.1	14.8

^aSource: Medicare Current Beneficiary Survey (MCBS) 2001, 2003, and 2005

^bNo statistically significant differences between the estimation and validation sample were observed at a .05 alpha level

^cADL= activities of daily living, IADL=instrumental activities of daily living, difficulty determined as requiring assistance from other individuals, use of equipment, or can't perform activity

Frequency of Claims Based Indicators, Overall and by Survey-Based Disability Status in Estimation Half of Sample ^{a,b}

Table 2

Modified Berenson-Eggers Type of Service (BETOS) Procedures Groups	Full Estimation Sample (n=7,394)		Good DS (PS=0,1,2) (n=6,707)		Poor DS (PS=3,4) (n=687)	
	Frequency	%	Frequency	%	Frequency	%
<u>Evaluation and Management (E&M)/other visits</u>						
Nursing home visit	747	10.10	329	4.91	418	60.84
Dermatology E&M visit	922	12.47	887	13.22	35	5.09
Neurology E&M visit	339	4.58	293	4.37	46	6.70
Rheumatology E&M visit	180	2.43	165	2.46	15	2.18
Chiropractic ^c	490	6.63	-	-	-	-
Home visit	121	1.64	76	1.13	45	6.55
Hospice visit	68	0.92	23	0.34	45	6.55
<u>Minor procedures</u>						
Skin	1740	23.53	1,394	20.78	346	50.36
<u>Ambulatory procedures</u>						
Musculoskeletal	125	1.69	114	1.70	11	1.60
<u>Preventive services</u>						
Screenings	2672	36.14	2,610	38.91	62	9.02
Immunizations/vaccinations	3384	45.77	3,236	48.25	148	21.54
<u>Major procedures</u>						
Cardiovascular	423	5.72	383	5.71	40	5.82
Colectomy ^c	36	0.49	-	-	-	-
Orthopedic - other	130	1.76	101	1.51	29	4.22
<u>Durable Medical Equipment</u>						
Bath and toilet aids	140	1.89	107	1.60	33	4.80
Wheelchairs	365	4.94	231	3.44	134	19.51
Hospital bed	173	2.34	74	1.10	99	14.41
Enteral and parenteral	50	0.68	11	0.16	39	5.68

Modified Berenson-Eggers Type of Service (BETOS) Procedures Groups	Full Estimation Sample		Good DS (PS=0,1,2)		Poor DS (PS=3,4)	
	(n=7,394)		(n=6,707)		(n=687)	
	Frequency	%	Frequency	%	Frequency	%
Medical/surgical supplies	186	2.52	149	2.22	37	5.39
Other	917	12.40	787	11.73	130	18.92
Imaging						
Imaging/procedure - heart including cardiac catheter	249	3.37	237	3.53	12	1.75
Standard imaging - nuclear medicine	936	12.66	873	13.02	63	9.17
Other						
Ambulance	983	13.29	675	10.06	308	44.83
Electrocardiography monitoring & cardiovascular stress tests	960	12.98	916	13.66	44	6.40
Endoscopy - upper gastrointestinal	371	5.02	326	4.86	45	6.55
Endoscopy - sigmoidoscopy, colonoscopy	636	8.60	601	8.96	35	5.09

^aSource: Medicare Current Beneficiary Survey (MCBS) 2001, 2003, and 2005

^bSignificant (P-value <.05) rows highlighted in dark gray indicate higher frequency among those with poor relative to good DS; light gray indicates a greater proportion good DS compared to poor DS.

^cRows hid due to small cell sizes (<11 individuals)

Predictive Validity of Disability Status Models with and without Medicaid and Region Interactions^{a,b}

Table 3

	Survey-Based DS				Negative Predictive Value	Positive Predictive Value
	Good (DS = 0,1,2)	Poor (DS = 3,4)	Sensitivity	Specificity		
Within estimation sample						
Model without interactions	90.7	90.3	80.9	91.8	97.9	50.3
Model with interactions			81.2	91.6	97.9	49.7
Within validation sample						
Model without interactions	91.1	9.0	79.0	91.7	97.8	48.3
Model with interactions			78.0	91.6	97.7	47.7
Within validation sample by self report health status						
Model without interactions						
Poor	55.3	44.7	78.8	76.5	84.2	69.3
Fair, good, very good, or excellent	91.9	8.1	79.1	92.4	98.4	42.8
Model with interactions						
Poor	59.7	40.3	76.3	77.1	82.8	69.3
Fair, good, very good, or excellent	93.3	6.7	78.5	92.6	98.4	43.4

^aSource: Medicare Current Beneficiary Survey (MCBS) 2001, 2003, and 2005

^bSelf report health status poor n=491 and fair, good, very good, excellent n=6,903, prediction sample n=7,394 validation sample n=7,394

Table 4
Effects of Survey-Based and Claims-Based Disability Status Measures on Survival in Medicare Beneficiaries^{a,b}

	Model 1			Model 2			Model 3			Model 4		
	No CCI or DS			Survey-based DS only			Claims-based DS only			Both claims-based DS and CCI		
	HR	CI	95% HR	HR	CI	95% HR	HR	CI	95% HR	HR	CI	95% HR
Age												
65-69 (ref)	-	-	-	-	-	-	-	-	-	-	-	-
70-74	1.21	0.98	1.48	1.18	0.96	1.45	1.25	1.02	1.53	1.24	1.01	1.52
75-79	1.77	1.45	2.15	1.76	1.44	2.14	1.79	1.47	2.18	1.71	1.41	2.09
80-84	2.57	2.13	3.11	2.36	1.95	2.85	2.44	2.02	2.95	2.32	1.92	2.80
85+	5.49	4.58	6.51	4.37	3.63	5.26	4.44	3.69	5.33	4.23	3.52	5.09
Race												
White (ref)	-	-	-	-	-	-	-	-	-	-	-	-
Non-white	1.01	0.88	1.16	1.03	0.90	1.18	1.18	1.03	1.36	1.17	1.02	1.34
Sex												
Female (ref)	-	-	-	-	-	-	-	-	-	-	-	-
Male	0.67	0.61	0.75	0.63	0.57	0.70	0.63	0.57	0.70	0.66	0.60	0.73
Educational attainment												
< HS	1.42	1.27	1.58	1.24	1.11	1.39	1.25	1.12	1.40	1.19	1.06	1.33
HS graduate	1.12	0.99	1.26	1.09	0.97	1.23	1.11	0.99	1.26	1.09	0.97	1.23
> HS graduate(ref)	-	-	-	-	-	-	-	-	-	-	-	-
Marital status												
Married (ref)	-	-	-	-	-	-	-	-	-	-	-	-
Not married	1.52	1.37	1.69	1.50	1.35	1.67	1.32	1.19	1.47	1.31	1.17	1.45
Household income												
< \$25,000	0.88	0.70	1.10	0.81	0.64	1.01	0.77	0.62	0.97	0.80	0.64	1.00
\$25,001-\$50,000	1.12	0.85	1.46	1.05	0.80	1.37	1.06	0.81	1.39	1.06	0.81	1.38
> \$50,000 (ref)	-	-	-	-	-	-	-	-	-	-	-	-
Residence location												

	Model 1			Model 2			Model 3			Model 4		
	No CCI or DS			Survey-based DS only			Claims-based DS only			Both claims-based DS and CCI		
	HR	CI	95% HR	HR	CI	95% HR	HR	CI	95% HR	HR	CI	95% HR
Rural	1.01	0.92	1.11	1.01	0.91	1.11	1.01	0.92	1.11	1.04	0.94	1.14
Urban (ref)	-	-	-	-	-	-	-	-	-	-	-	-
Survey based DS												
0,1,2 (ref)	-	-	-	-	-	-	-	-	-	-	-	-
3,4	-	-	-	3.97	3.55	4.44	-	-	-	-	-	-
Claims based DS												
0,1,2 (ref)	-	-	-	-	-	-	-	-	-	-	-	-
3,4	-	-	-	-	-	-	3.53	3.18	3.92	2.85	2.56	3.17
CCI												
0 (ref)	-	-	-	-	-	-	-	-	-	-	-	-
1	-	-	-	-	-	-	-	-	-	1.37	1.21	1.55
2+	-	-	-	-	-	-	-	-	-	2.16	1.93	2.42

^aSource: Medicare Current Beneficiary Survey (MCBS) 2001, 2003, and 2005

^bWithin validation half of the sample (n=7,394), observation time is from date of fall MCBS survey in 2001, 2003, or 2005 until death or December 31st, 2007