

# Comorbidity Technical Report

## The Impact of Different SEER-Medicare Claims-based Comorbidity Indexes on Predicting Non-cancer Mortality for Cancer Patients

### Authors

Margaret R. Stedman, PhD<sup>1,2</sup>; Paul Doria-Rose, PhD<sup>2</sup>; Joan L. Warren, PhD<sup>2</sup>; Carrie N. Klabunde, PhD<sup>3</sup>; Angela Mariotto, PhD<sup>2</sup>

<sup>1</sup> Stanford University School of Medicine, Palo Alto, CA

<sup>2</sup> Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

<sup>3</sup> Office of Disease Prevention, Office of the Director, National Institutes of Health, Bethesda, MD

### For More Information

Angela Mariotto, PhD  
[mariotta@mail.nih.gov](mailto:mariotta@mail.nih.gov)

### Abstract

**Background:** Comorbidity indexes are useful summaries of overall health. These indexes can be used to predict risk of mortality and to reduce confounding in observational studies. Although comorbidity indexes and condition weights may vary, it is not clear what impact use of different weights may have on predicting non-cancer deaths for cancer patients. The purpose of this report is to evaluate variations in predicting non-cancer death for cancer patients using differing weights applied to the conditions included in the original Charlson index.

**Methods:** We used the 2014 SEER-Medicare linked database, to define five cohorts of cancer patients diagnosed between 1992 and 2009: all cancer sites (n=1,618,875), breast (n=204,277), prostate (n=293,094), colorectal (n=185,374), and lung (n=244,068). We identified 16 comorbid conditions using a revised list of ICD-9 codes found on Medicare claims. We estimated weights using a Cox proportional hazard model for deaths due to causes other than cancer for each of the 5 cohorts. The analysis included an additional set of weights for the all cancer cohort which allowed for two-way interactions between different comorbidity conditions. For each patient we calculate their comorbidity index using the sum of their comorbidity weights. We evaluated each index by assessing the accuracy of predicting non-cancer related death at 1, 3, and 5 years post incident cancer diagnosis using time-dependent ROC statistics for different cohorts of cancer patients.

**Results:** Site-specific indexes had similar accuracy to indexes based on all cancer sites or the Charlson index. The index containing weights for 2-way interactions of combinations of comorbid conditions had slightly improved predictive accuracy over indexes with weights for main effects only. Predictive accuracy was 70%-80% for most models and cohorts tested.

**Conclusions:** There were only slight differences in the predictive accuracy of the indexes tested. The weights from the original Charlson index had similar accuracy to weights estimated for the NCI Index. Thus, differences in weights for the same set of conditions appear to have small effects on non-cancer mortality predictions. These indexes do not include other factors, such as smoking history and socio-economic status, which may better predict mortality beyond the 16 Charlson conditions.

## **Introduction**

Comorbidity indexes are commonly used in analyses to measure and adjust for disease burden among patients in health services research studies. Composite indexes provide a sum, often weighted, of a select number of chronic health conditions that have been found to be associated with patient survival. One of the most well-known and widely used methods to measure comorbidity is the Charlson Index<sup>1</sup> which was developed in 1987 to predict 1 year mortality from hospitalized patients. Charlson et al. assigned empirically derived weights to 19 medical conditions,

and summed them into a composite measure of illness burden. The Charlson index was later adapted for use with healthcare claims data by Deyo<sup>2</sup> and Romano<sup>3</sup>.

In 2000, the NCI Comorbidity Index was developed by Klabunde and colleagues to predict non-cancer related deaths in cancer patients<sup>4</sup> using the SEER-Medicare claims data. The NCI Index included 16 of the 19 conditions in the Charlson index; because of the focus on studying comorbid conditions in cancer patients, solid cancers, leukemias and lymphomas were excluded. ICD-9 codes identified by Richard Deyo<sup>2</sup> and Patrick Romano<sup>3</sup> and CPT-4 codes<sup>5</sup> were used to identify comorbid conditions from claims. The NCI Index includes disease conditions ascertained from both inpatient and physician claims for a one year period prior to the month of diagnosis. Site-specific weights were developed for breast, prostate, colorectal, and lung cancer patients. In 2013, Mariotto et al. developed comorbidity adjusted lifetables constructed from the comorbid conditions and weights that included interactions between conditions included in the NCI Index.<sup>6</sup>

The diagnosis codes used to assess comorbidity have not been updated since their initial development. In an effort to update the NCI Comorbidity Index, staff at the NCI undertook a detailed review of the codes and algorithms used to [identify comorbidities](#). The new codes used to identify the conditions have been translated into new SAS macros. The purpose of this report is to use the updated codes to calculate different comorbidity indexes based on different weights/indexes and to compare their ability to predict non-cancer death for specific cohorts of cancer patients. We focused on weights for: specific cancer sites, ii) all cancer sites combined, and iii) the original Charlson index.

## **Methods**

### **Data Source and Population**

The NCI Comorbidity Index was developed for use with claims from the SEER-Medicare Data<sup>7</sup>, a database linking the Surveillance Epidemiology and End Results (SEER) data with Medicare claims. The SEER database originates from 18 cancer registries located in 9 states and 6 metropolitan areas, representing 28% of the US population.<sup>8</sup> The SEER

data contain information about the occurrence and characteristics of primary incident cancers, including diagnosis date, stage, grade, initial treatment, and cause of death information.

The linkage with Medicare claims began in 1992 and is updated biannually, with the most recent 2016 linkage update including cancer patients diagnosed in 2013. For persons in the SEER data who are age 65 or older, 93% are linked to Medicare enrollment records.<sup>7</sup> The Medicare data used for this assessment included hospital and physician claims for beneficiaries who had both Medicare Part A and Part B services.

This analysis involved patients from the SEER-Medicare data who were ages 66 and older with a cancer diagnosis between 1992 and 2011. The analysis included only those patients residing in the first 11 cancer registries to join the SEER program. These 5 states and 6 metropolitan areas represent 14% of the US population. Patients were followed for survival outcomes (cancer or non-cancer related death) from 1992 until December 31, 2011. Patients were excluded if their cancer was diagnosed at the time of death as noted on death certificate or autopsy. Only cases of primary tumors were selected for analysis.

Five cohorts were defined based on the primary cancer site at diagnosis. These included: breast (ICD-O-3 sites: C500-C509), prostate (ICD-O-3 site: C619), lung and bronchus (ICD-O-3 sites: C340-C349), colorectal (ICD-O-3 sites: C180, C182-C187, C199, C209), and all cancer sites combined. Comorbid conditions were ascertained in the year prior to diagnosis excluding month of diagnosis to avoid capturing sequelae of treatment. Conditions identified from the codes in the physician claims but not in hospital claims were required to appear more than once in a period greater than 30 days<sup>4</sup>. This requirement aimed to exclude conditions that were “ruled out.”

### **Codes Used to Define Comorbid Conditions**

NCI worked with professional coders from [Care Communications Inc](#)<sup>8</sup> to review the diagnosis and procedure codes used to define each of the conditions included in the NCI Comorbidity Index. Coders from Care Communications Inc. reviewed all original codes to confirm that the codes were accurate, and relevant to the conditions and procedures listed in the original macro. They also identified new codes that might be used to identify conditions.

All suggested new codes were reviewed by co-authors Warren and Doria-Rose to produce the final set of definitions. We grouped Diabetes and Diabetes + complications into a single group Any Diabetes, and Mild Liver Disease and Mod/Severe Liver Disease into Any Liver Disease to create 14 categories of conditions from the 16 conditions originally identified.

### **Comorbidity Weights**

Weights were estimated from Cox proportional hazard models (model based weights) predicting non-cancer related death. We used the SEER cause-specific death classification algorithm to define non cancer related death.<sup>9</sup> Patients were censored at time of cancer death or end of follow-up. The models included the comorbid conditions significantly associated with non-cancer related death after adjusting for age, gender, and race. The weights are the estimated coefficients from the Cox proportional hazards models.

We estimated the site-specific weights using respective data for breast, colorectal, lung and prostate cancer cohorts. For all sites combined we estimated 2 sets of weights -- one similar to the site-specific weights with age, race, and gender and significant comorbid conditions and additional weights including two-way interaction for the most prevalent conditions. This is consistent with the approach used by Mariotto et al.<sup>6</sup>. In total, 6 different model-based weighting schemes were included in this analysis: four for the site-specific cohorts with main effects alone and 2 weighting schemes (main effects alone and main effects plus interactions) for the all sites combined cohort.

Once the weights were developed, an individualized comorbidity index, CI, could be calculated from the sum of the weights specific to the comorbid conditions of the patient. For example, a patient with diabetes and COPD, would have a comorbidity index that is the sum of the weights (model coefficients) associated with diabetes and COPD. In contrast, a patient with none of the 16 relevant conditions would have a comorbidity index of zero.

## Predictive Accuracy of Comorbidity Indexes

The comorbidity indexes were compared on the basis of predictive accuracy and model fit. Each site-specific index was compared to the Charlson weights and to weights estimated from the all sites combined cohort. Each site-specific index was tested in its respective site-specific cohort except for the Charlson and the all sites combined indexes which were tested in all the cohorts. Model accuracy was assessed in each of the 5 cohorts using the time dependent area under the curve (AUC(t)) at years 1, 3, and 5 from diagnosis<sup>10</sup>. The riskset ROC package from R was used to estimate AUC(t). Models were also evaluated for over-fitting using the shrinkage estimator recommended by Harrell et al.<sup>11</sup>

## Results

In total, 1,618,875 patients were included in the analysis (Table 1). Lung cancer patients had the greatest number of comorbidities (13.3% with 3+ comorbidities); 37.5% of lung cancer patients had coexisting chronic pulmonary disease. In contrast, breast and prostate cancer patients were healthier, with 61.2% and 63.8%, respectively, having no comorbid conditions. The colorectal cancer group was older (19.0% age 85+) and had a moderate comorbidity burden (9.0% of CRC patients had 3+ comorbidities).

As the original Charlson weights were based on the hazard ratio rounded to the nearest integer, we present the hazard ratios from the new models alongside the original Charlson weights. (Table 2) Many of the conditions had hazard ratios similar to the original Charlson weights, however there are a few noteworthy differences. In the original Charlson index, congestive heart failure had a weight of 1 compared with site specific Cox HRs that ranged from 1.83 (lung) to 1.90 (breast). Dementia also had a Charlson weight of 1 compared with site specific Cox HRs of 1.91 (lung) to 2.29 (prostate), about double the risk of non-cancer related death. In contrast, in the original Charlson index, an AIDS diagnosis had a weight of 6 but had site specific Cox HRs for other cause death ranging from 1.67 (prostate) to 2.36 (lung). There was little variability in the estimated weights across the site specific cohorts and all interaction term coefficients were negative in the all cancer sites model. The actual weights, based on the parameter estimates, included in the comorbidity index are listed in Table 3.

Predictive accuracy of the models was similar across models and cohorts (Table 4). Under ideal conditions predictive accuracy is 100% ; models with accuracy below 50% are considered to be poor predictors of non-cancer related death. We found all models to be reasonably accurate with AUC(t)s above 0.5. Predictive accuracy was highest in the first year after diagnosis and declined steadily 3 and 5 years post cancer diagnosis. On average, the index based on interactions tended to have accuracy comparable to that of the other models (all sites 0.724 vs. 0.721) and to the original Charlson index (all sites=0.717). However, from previous study they tend to better predict mortality for other causes for the small group of patients with more than 1 comorbidity.<sup>5</sup> There was no evidence of over-fitting for any of the models (shrinkage factor was approximately 1, results not shown).

## **Discussion**

These findings show that the newly revised weights estimated from the Cox models offer slight improvement in predicting non-cancer related death compared to the original Charlson condition weights, with the greatest difference in prediction in the first year after diagnosis. Weighted indexes with interactions had slightly higher predictive accuracy, while in cancer site-specific cohorts indexes with main effects alone had similar accuracy regardless of whether all sites combined or site-specific weights were used. Either type of index may be selected depending on the characteristics of the study cohort.

Studies involving populations with a greater burden of comorbidity may benefit from the new index with interactions, although the impact would be modest. An index with main effects alone would likely overestimate the risk of non-cancer related death for patients with more comorbidity and reduce the predictive accuracy of the model. In this analysis 20% of the patient population had at least two comorbid condition. Predictive accuracy of the indexes improved slightly for all cancer sites, main effects (0.683) to all cancer sites, main effects + interactions (0.688) at five years post diagnosis.

Site-specific indexes were comparable to indexes derived from all cancer sites. However, for the lung cancer cohort, we found predictive accuracy was slightly worse for the site-specific index, particularly for predictions years 3 and 5 after diagnosis. Some of the loss in accuracy may be related to fewer non-cancer related deaths in the

cohort. The lung cancer cohort had only 13% non-cancer related deaths compared with the other cohorts where the percentage of non-cancer related deaths was above 25%. For sites where there are few non-cancer related deaths the index derived from all cancer-sites may offer more stability than the site-specific index.

The Charlson index was originally derived from data from the early 1980's; our study included data from 1992-2011 so that some shift in the weights assigned to comorbid conditions may be influenced by recent advances in medical technology and increases in life expectancy. We observed a shift toward lower weights for AIDS, and higher weights for dementia and cardiovascular disease in our models. Improved survival in patients with AIDS can be explained by the introduction of highly active antiretroviral therapy in 1997, which has dramatically improved outcomes for HIV infected patients. Other differences in weights may be explained by inherent differences in the SEER-Medicare cancer patient population compared to the Charlson population of "all patients admitted to the medical service at New York Hospital"<sup>1</sup>. For other conditions, there were few changes. For example, liver disease and renal disease are important predictors of mortality in both the original Charlson index and the NCI Comorbidity index.

Our methods relied on the SEER cause-specific death classification<sup>2</sup> to determine whether the cause of death was cancer related or due to other causes. This definition was used to improve the causes of death information from death certificates. This could result in comorbid condition weights that are slightly attenuated towards the null, as was observed in the lung cancer cohort. Additionally, comorbid conditions were ascertained from administrative claims data, which are not as accurate as medical chart abstraction.<sup>12</sup> Medicare claims only allow for a specific number diagnoses to be recorded so we may be missing diagnoses that could have contributed to the patient's overall score. In addition, the sensitivity and specificity of our new definitions has not been determined. Another limitation is the fact that we restrict conditions to the conditions examined by Charlson in the 1980s. We do not include hypertension or other conditions that may affect mortality.

Since updated definitions and condition weights only yielded slight improvements in predictive accuracy, it may be more important to consider other factors that contribute to the risk of non-cancer related death. Sharabiani et al.<sup>13</sup> conducted a systematic review to determine the best performing comorbidity measures for mortality. They found the Elixhauser comorbidity index offered a slightly better prediction of long term mortality compared with



the Charlson index. The Elixhauser AHRQ-Web ICD-9 coding algorithm includes 30 comorbid conditions, 6 of which overlap with the Charlson index. Conditions that are unique to the Elixhauser index include depression, alcohol misuse, drug abuse, obesity, weight loss, psychoses, other neurological disorders, anemia, hypothyroidism, and hypertension.<sup>14</sup> The original Elixhauser coding algorithm did not attempt to summarize comorbidities into a single score; however van Walraven et al.<sup>15</sup> later modified the Elixhauser algorithm to a point system to create a comorbidity risk score. Comorbidities with the highest points for hospital death included cancer, liver disease, paralysis, and congestive heart failure. Conditions such as anemia, obesity, drug abuse, and depression had a negative effect on the overall score. Nine of the 30 conditions, including AIDS, did not contribute to the overall score. Although less detailed than the Elixhauser method, our condition weights correspond with the points assigned to the Elixhauser disease risk score. It is possible that including some of the Elixhauser conditions may have a greater impact on the predictive accuracy of the score than revising the weights included in the score. However prior work comparing Elixhauser, Charlson, and the NCI comorbidity index did not demonstrate the Elixhauser algorithm to perform substantially better in a cancer patient population.<sup>16</sup>

The updated NCI Comorbidity Indexes are comparable to the initial index<sup>4</sup> in predicting non-cancer related mortality in cancer patients. Users of the NCI Comorbidity Index can determine the best algorithm, site-specific or scores with interactions, depending on the characteristics of their study cohort. Disease and procedure codes evolve over time and it is important to re-evaluate and update definitions to identify conditions and condition weights to reflect the current state of medical care. We anticipate the transition from ICD-9-CM to ICD-10-CM will offer new challenges and require adjustments to our current algorithm.

Table 1: Characteristics of study population of SEER-Medicare patients diagnosed between 1992-2011 by cancer site.

<b>Demographics</b>	<b>All Sites N=1,618,875</b>	<b>Breast N=204,277</b>	<b>Prostate N=293,094</b>	<b>Lung N=244,068</b>	<b>CRC N= 185,374</b>
<b>Age at diagnosis</b>					
<b>66-74</b>	735,052 (45.4%)	96,251 (47.1%)	164,901 (56.3%)	114,656 (47.0%)	70,524 (38.0%)
<b>75-84</b>	658,595 (40.7%)	81,337 (39.8%)	106,962 (36.5%)	102,413 (42.0%)	79,656 (43.0%)
<b>85+</b>	225,228 (13.9%)	26,689 (13.1%)	21,231 (7.2%)	26,999 (11.1%)	35,194 (19.0%)
<b>Race</b>					
<b>White</b>	1,408,361 (87.0%)	180,150 (88.2%)	245,515 (83.8%)	212,269 (87.0%)	159,135 (85.9%)
<b>Black</b>	135,197 (8.4%)	16,069 (7.9%)	33,750 (11.5%)	20,861 (8.6%)	16,437 (8.9%)
<b>Other</b>	75,317 (4.7%)	8,058 (3.9%)	13,829 (4.7%)	10,938 (4.5%)	9,802 (5.3%)
<b>Sex</b>					
<b>Male</b>	849,387 (52.5%)	0 (0%)	293,094 (100.0%)	128,534 (52.7%)	85,053 (45.9%)
<b>Female</b>	769,488 (47.5%)	204,277 (100.0%)	0 (0%)	115,534 (47.3%)	100,321 (54.1%)
<b>Number of Cancer deaths</b>	618,896 (38.2%)	29,474 (14.4%)	34,534 (11.8%)	181,458(74.4%)	63,480 (34.2%)
<b>Number of Other-Cause Deaths</b>	404,677 (25.0%)	57,893 (28.3%)	95,008 (32.4%)	34,260 (14.0%)	56,798 (30.6%)
<b>Number Alive</b>	595,302 (36.8%)	87,367 (42.8%)	163,552 (55.8%)	28,350 (11.6%)	65,096 (35.1%)
<b>Number of Comorbid conditions</b>					
<b>0</b>	854,771 (52.8%)	125,108 (61.2%)	187,073 (63.8%)	96,913 (39.7%)	97,354 (52.5%)
<b>1</b>	437,498 (27.0%)	51,630 (25.3%)	69,974 (23.9%)	75,489 (30.9%)	49,639 (26.8%)
<b>2</b>	187,146 (11.6%)	17,432 (8.5%)	22,900 (7.8%)	39,099 (16.0%)	21,628 (11.7%)
<b>3+</b>	139,460 (8.6%)	10,107 (5.0%)	13,147 (4.5%)	32,567 (13.3%)	16,753 (9.0%)
<b>Comorbid Conditions (based on updated definition):</b>					
<b>Acute MI</b>	22,208 (1.4%)	1,532 (0.8%)	2,879 (0.98%)	4,333 (1.78%)	3,313 (1.8%)
<b>History of MI</b>	40,193 (2.5%)	2,490 (1.2%)	5,987 (2.04%)	8,154 (3.34%)	4,753 (2.6%)
<b>CHF</b>	175,357 (10.8%)	14,923 (7.3%)	18,757 (6.40%)	34,005 (13.93%)	23,756 (12.8%)
<b>PVD</b>	144,892 (9.0%)	13,399 (6.6%)	16,453 (5.61%)	31,423 (12.87%)	16,833 (9.1%)
<b>CVD</b>	113,320 (7.0%)	10,748 (5.3%)	14,914 (5.09%)	21,162 (8.67%)	13,846 (7.5%)
<b>Chronic Pulmonary</b>	276,372 (17.1%)	21,770 (10.7%)	29,925 (10.21%)	91,524 (37.50%)	26,689 (14.4%)

**Table 1: Characteristics of study population of SEER-Medicare patients diagnosed between 1992-2011 by cancer site.**

<b>Demographics</b>	<b>All Sites N=1,618,875</b>	<b>Breast N=204,277</b>	<b>Prostate N=293,094</b>	<b>Lung N=244,068</b>	<b>CRC N= 185,374</b>
<b>Dementia</b>	47,568 (2.9%)	5,281 (2.6%)	4,302 (1.47%)	7,398 (3.03%)	6,767 (3.6%)
<b>Paralysis</b>	13,569 (0.8%)	1,209 (0.6%)	1,810 (0.62%)	2,178 (0.89%)	1,819 (1.0%)
<b>Any Diabetes</b>	328,617 (20.3%)	37,588 (18.4%)	49,672 (16.95%)	46,928 (19.23%)	39,602 (21.4%)
<b>Renal</b>	74,145 (4.6%)	5,437 (2.7%)	9,571 (3.27%)	11,990 (4.91%)	7,810 (4.2%)
<b>Any Liver Disease</b>	14,751 (0.9%)	909 (0.4%)	999 (0.34%)	1,710 (0.70%)	1,094 (0.6%)
<b>Peptic Ulcer</b>	26,406 (1.6%)	1,834 (0.9%)	2,902 (0.99%)	4,254 (1.74%)	3,580 (1.9%)
<b>Rheumatologic</b>	37,342 (2.3%)	5,025 (2.5%)	3,556 (1.21%)	7,059 (2.89%)	3,762 (2.0%)
<b>AIDS</b>	723 (0.0%)	22 (0.0%)	136 (0.1%)	129 (0.1%)	48 (0.0%)

MI=myocardial infarction

Table 2: Hazard Ratios from modeling other-cause death associated with the comorbid conditions adjusting for age, sex, and race

Comorbid Condition	All Cancer Sites Charlson <sup>1</sup> (original condition weight)	Hazard Ratios from Cox models of other-cause death adjusting for age, sex, race					
		All Cancer Sites		Breast	Prostate	Lung	CRC
		Main	Main + interac	Main	Main	Main	Main
Acute MI	1	1.14	1.15	NS	1.10	1.30	1.11
History of MI	1	1.08	1.13	1.15	1.14	NS	NS
CHF	1	1.91	2.15	1.90	1.97	1.83	1.89
PVD	1	1.30	1.43	1.35	1.31	1.18	1.27
CVD	1	1.32	1.50	1.41	1.37	1.19	1.32
Chronic Pulmonary	1	1.69	1.91	1.63	1.71	1.40	1.57
Dementia	1	2.06	2.07	2.27	2.29	1.91	2.01
Paralysis	2	1.49	1.50	1.55	1.43	1.51	1.48
Diabetes	1						
Diabetes + complications	2						
Any Diabetes		1.34	1.41	1.49	1.36	1.14	1.37
Renal	2	1.60	1.77	1.95	1.54	1.46	1.69
Mild Liver Disease	1						
Mod/Severe Liver Dis	3						
Any Liver Disease		2.09	2.12	2.50	1.79	1.53	2.05
Peptic Ulcer	1	1.08	1.17	1.10	NS	1.09	NS
Rheumatologic	1	1.25	1.35	1.25	1.16	1.18	1.25
AIDS	6	1.79	1.76	NS	1.67	2.36	NS
CHF* Rheumatologic			0.95				
CHF*Old MI			0.93				
CHF*Renal			0.96				
CHF*PVD			0.86				

Table 2: Hazard Ratios from modeling other-cause death associated with the comorbid conditions adjusting for age, sex, and race

Comorbid Condition	All Cancer Sites Charlson <sup>1</sup> (original condition weight)	Hazard Ratios from Cox models of other-cause death adjusting for age, sex, race					
		All Cancer Sites		Breast	Prostate	Lung	CRC
		Main	Main + interac	Main	Main	Main	Main
CHF*CVD			0.81				
CHF*Chronic Pulmonary			0.90				
CHF*Diabetes			NS				
CHF*Ulcer			0.90				
CHF*Acute MI			NS				
Chronic Pulmon*Rheumatologic			0.91				
Chronic Pulmon*Old MI			NS				
Chronic Pulmon*Renal			0.85				
Chronic Pulmon*PVD			0.92				
Chronic Pulmon*CVD			0.84				
Chronic Pulmon*Diabetes			0.86				
Chronic Pulmon*Ulcer			0.94				
Chronic Pulmon*Acute MI			NS				
Diabetes*Rheumatologic			0.85				
Diabetes*Old MI			NS				
Diabetes*Renal			NS				
Diabetes *PVD			0.95				
Diabetes *CVD			0.96				
Diabetes *Ulcer			0.92				
Diabetes *Acute MI			1.15				

NS: Not significant parameters were dropped from the final model

CHF=congestive heart failure; MI=myocardial infarction; CVD=cerebrovascular disease; PVD=peripheral vascular disease

Table 3: Comorbid condition weights obtained from modeling other-cause death associated with the conditions adjusting for age, sex, and race

Comorbid Condition	Charlson <sup>1</sup> condition weight (original)	Parameter estimates from Cox models of other-cause death adjusting for age, sex, race					
		All Cancer Sites		Breast	Prostate	Lung	CRC
		Main	Main + interac	Main	Main	Main	Main
Acute MI	1	0.12624	0.14234	NS	0.095	0.26543	0.1043
History of MI	1	0.07999	0.12617	0.13951	0.1295	NS	NS
CHF	1	0.64441	0.76496	0.64212	0.6789	0.60435	0.6357
PVD	1	0.26232	0.35867	0.29706	0.2697	0.16764	0.2364
CVD	1	0.27868	0.40263	0.3463	0.3131	0.17367	0.2783
Chronic Pulmonary	1	0.52487	0.64533	0.48774	0.5383	0.335	0.4535
Dementia	1	0.72219	0.72943	0.81842	0.8279	0.64843	0.6974
Paralysis	2	0.39882	0.40474	0.43507	0.3573	0.40968	0.393
Diabetes	1						
Diabetes + complications	2						
Any Diabetes		0.29408	0.34405	0.40134	0.3095	0.13148	0.3167
Renal	2	0.4701	0.57191	0.66631	0.431	0.37908	0.5232
Mild Liver Disease	1						
Mod/Severe Liver Dis	3						
Any Liver Disease		0.73803	0.75217	0.91659	0.5792	0.42426	0.7192
Peptic Ulcer	1	0.07506	0.15717	0.09235	NS	0.09023	NS
Rheumatologic	1	0.21905	0.29964	0.22226	0.1512	0.16457	0.2216
AIDS	6	0.58362	0.56621	NS	0.5127	0.85973	NS
CHF* Rheumatologic			-0.05459				
CHF*Old MI			-0.07711				
CHF*Renal			-0.04482				
CHF*PVD			-0.15692				

Table 3: Comorbid condition weights obtained from modeling other-cause death associated with the conditions adjusting for age, sex, and race

Comorbid Condition	Charlson <sup>1</sup> condition weight (original)	Parameter estimates from Cox models of other-cause death adjusting for age, sex, race					
		All Cancer Sites		Breast	Prostate	Lung	CRC
		Main	Main + interac	Main	Main	Main	Main
CHF*CVD			-0.20523				
CHF*Chronic Pulmonary			-0.11012				
CHF*Diabetes			NS				
CHF*Ulcer			-0.10874				
CHF*Acute MI			NS				
Chronic Pulmon*Rheumatologic			-0.09531				
Chronic Pulmon*Old MI			NS				
Chronic Pulmon*Renal			-0.16777				
Chronic Pulmon*PVD			-0.08479				
Chronic Pulmon*CVD			-0.17704				
Chronic Pulmon*Diabetes			-0.15243				
Chronic Pulmon*Ulcer			-0.06183				
Chronic Pulmon*Acute MI			NS				
Diabetes*Rheumatologic			-0.16149				
Diabetes*Old MI			NS				
Diabetes*Renal			NS				
Diabetes *PVD			-0.05465				
Diabetes *CVD			-0.03712				
Diabetes *Ulcer			-0.07896				
Diabetes *Acute MI			NS				

CHF=congestive heart failure; MI=myocardial infarction; CVD=cerebrovascular disease; PVD=peripheral vascular disease

Table 4: Model accuracy in predicting other-cause survival adjusting for age, sex, and race measured by AUC(t) for t=1, 3, 5 and 10 years.

Cohort	Time	Charlson Index	All Cancer Sites		Site-Specific
			NCI Comorbidity Index, Main effects only	NCI Comorbidity Index, Main effects + interactions	NCI Comorbidity Index, Main Effects
All cancer sites	1-year	0.717	0.721	0.724	
	3-year	0.698	0.701	0.705	
	5-year	0.682	0.683	0.688	
	10-year	0.648	0.647	0.652	
CRC	1-year	0.714	0.719	0.722	0.719
	3-year	0.699	0.702	0.706	0.702
	5-year	0.686	0.687	0.691	0.687
	10-year	0.654	0.652	0.657	0.652
Breast	1-year	0.753	0.757	0.757	0.758
	3-year	0.734	0.735	0.737	0.736
	5-year	0.716	0.716	0.719	0.716
	10-year	0.682	0.679	0.682	0.679
Prostate	1-year	0.724	0.729	0.730	0.729
	3-year	0.707	0.709	0.712	0.709
	5-year	0.691	0.692	0.695	0.692
	10-year	0.655	0.654	0.657	0.654
Lung	1-year	0.645	0.652	0.655	0.650
	3-year	0.633	0.639	0.643	0.636
	5-year	0.621	0.627	0.633	0.623
	10-year	0.599	0.605	0.613	0.600



## References

1. 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(5):373-83. [\[View Abstract\]](#)
2. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992 Jun;45(6):613-9. [\[View Abstract\]](#)
3. 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 Oct;46(10):1075-9; discussion 1081-90. [\[Look up in PubMed\]](#)
4. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *J Clin Epidemiol* 2000 Dec;53(12):1258-67. [\[View Abstract\]](#)
5. American Medical Association. (2001) [Common Procedural Terminology- CPT 2002: Standard Edition.](#) Chicago, IL: AMA Press.
6. Mariotto AB, Wang Z, Klabunde CN, Cho H, Das B, Feuer EJ. Life tables adjusted for comorbidity more accurately estimate noncancer survival for recently diagnosed cancer patients. *J Clin Epidemiol* 2013 Dec;66(12):1376-85. doi: 10.1016/j.jclinepi.2013.07.002. [\[View Abstract\]](#)
7. National Cancer Institute. SEER-Medicare Data. <https://healthcaredelivery.cancer.gov/seermedicare/>. Accessed January 7, 2016
8. National Cancer Institute. Surveillance, Epidemiology, and End Results. <https://seer.cancer.gov registries/data.html>. Accessed January 7, 2016
9. Howlader N, Ries LA, Mariotto AB, Reichman ME, Ruhl J, Cronin KA. Improved estimates of cancer-specific survival rates from population-based data. *J Natl Cancer Inst* 2010 Oct 20;102(20):1584-98. doi: 10.1093/jnci/djq366. [\[View Abstract\]](#)
10. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics* 2005 Mar;61(1):92-105. [\[View Abstract\]](#)
11. Harrell FE Jr, Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. *J Natl Cancer Inst* 1988 Oct 5;80(15):1198-202. Review. [\[View Abstract\]](#)

12. Klabunde CN, Legler JM, Warren JL, Baldwin LM, Schrag D. A refined comorbidity measurement algorithm for claims-based studies of breast, prostate, colorectal, and lung cancer patients. *Ann Epidemiol* 2007 Aug;17(8):584-90. [\[View Abstract\]](#)
13. Sharabiani MT, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care* 2012 Dec;50(12):1109-18. Review. [\[View Abstract\]](#)
14. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998 Jan;36(1):8-27. [\[View Abstract\]](#)
15. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care* 2009 Jun;47(6):626-33. doi: 10.1097/MLR.0b013e31819432e5. [\[View Abstract\]](#)
16. 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 Aug;44(8):745-53. [\[View Abstract\]](#)