Predictive Performance of the Charlson Comorbidity Index: SNOMED CT Disease Hierarchy Versus International Classification of Diseases

Stephen P Fortin[1]
[1]Janssen R&D, LLC, Raritan, NJ, USA

Background

- Risk adjustment for disease severity and clinical prognosis is essential to obtaining valid inferences in causal inference research
- The Charlson Comorbidity Index (CCI) was originally developed as a weighted index comprised of 17 comorbid conditions to predict 1-year mortality risk among hospitalized patients
- In 2005, Quan et al. adapted the CCI, developing concept sets for each comorbid condition based on the International Classification of Diseases, Ninth and Tenth Revisions (ICD-9/10)
- Prior research has shown the use of the SNOMED CT Disease Hierarchy by the OHDSI Health Analytics Data-to-Evidence Suite's (HADES) FeatureExtraction package leads to a higher average CCI as compared to the Quan adaptation

Study Objectives: To compare the overlap in comorbid condition identification and predictive performance of the OHDSI (SNOMED CT) versus Quan (ICD) adaptations of the CCI.

Methods

Study Design: Descriptive study
Data Source: Data were from two U.S. administrative claims databases:

1. IBM® MarketScan® Multi-State Medicaid Database (MDCD)
2. Optum® De-Identified Clinformatics Data Mart Database – Date of Death (DOD)

Study Population: Patients aged ≥18 years with an inpatient visit between 01-01-2018 to 12-31-2018 with at least 365 days of prior observation (index = first inpatient visit)
Covariates: The CCI and each comorbid condition comprising the CCI was measured based on all observed diagnosis codes recorded at or any time prior to index. All covariates were measured using the OHDSI (SNOMED CT) versus Quan (ICD) adaptations of the CCI.
Statistical Analysis

- Descriptive statistics were produced for each study covariate
- For each comorbid condition, the overlap in patient capture between the SNOMED CT versus ICD vocabularies was described; and the average difference in the CCI attributable to each respective comorbid condition was calculated
- Logistic regression was used to develop a total of 5 models for 1-year mortality with following dependent variables:
  1. Age, and sex
  2. CCI (ICD)
  3. CCI (SNOMED CT)
  4. Age, sex, and CCI (ICD)
  5. Age, sex, and CCI (SNOMED CT)
- The predictive performance of each vocabulary was assessed using the c statistic, measured as the area under the curve of the receiver operating characteristics curve.

Results

- A total of 491,311 and 1,109,389 patients met the study criteria in MDCD and DOD, respectively
- SNOMED CT was associated with a higher average CCI as compared to ICD:

	MDCD	DOD
CCI, SNOMED CT vs. ICD	3.91 vs. 4.12	4.43 vs. 4.6

- A total of 24,017 (4.9%) and 145,516 (13.1%) deaths were observed in MDCD and DOD, respectively

Results

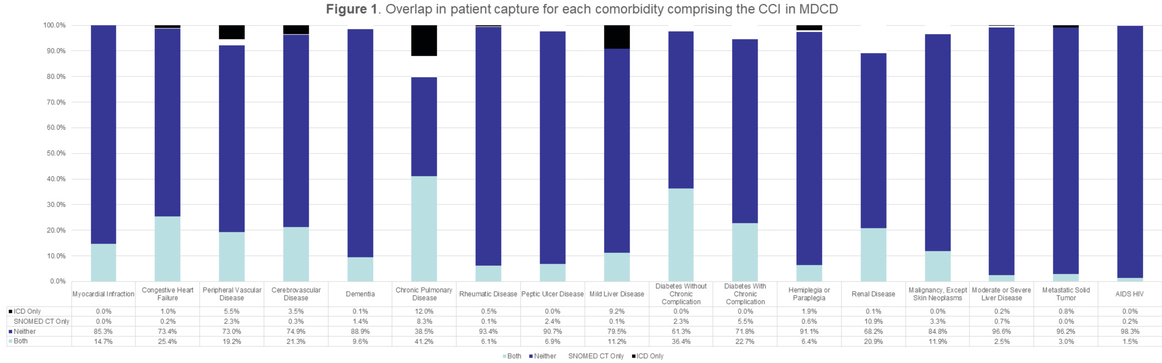
Figures 1 and 2 show the overlap in patient capture for each comorbidity in MDCD and DOD, respectively
Comorbid conditions identified in 5% of the study population by only SNOMED CT (and not ICD):

- Chronic pulmonary disease, diabetes with chronic complications, renal disease, malignancy

Comorbid conditions identified in 5% of the study population by only ICD (and not SNOMED CT):

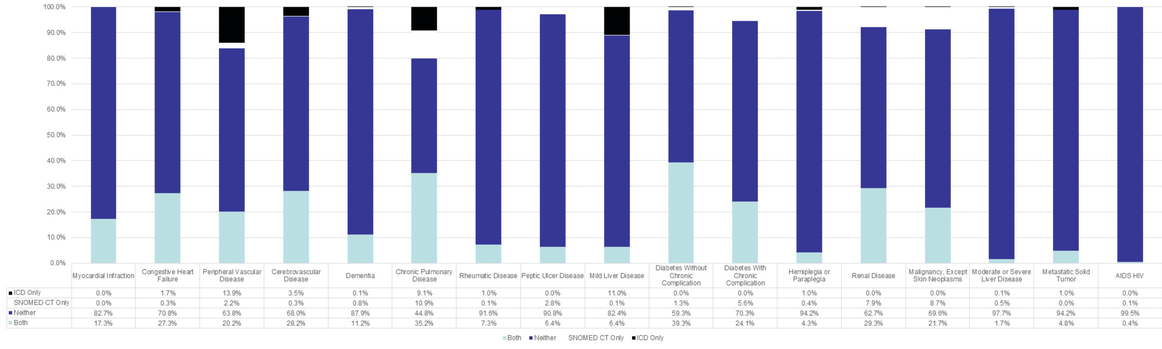
- Peripheral vascular disease, chronic pulmonary disease, mild liver disease

Figure 1. Overlap in patient capture for each comorbidity comprising the CCI in MDCD



	Myocardial Infraction	Congestive Heart Failure	Peripheral Vascular Disease	Cerebrovascular Disease	Dementia	Chronic Pulmonary Disease	Rheumatic Disease	Peptic Ulcer Disease	Mild Liver Disease	Diabetes Without Chronic Complication	Diabetes With Chronic Complication	Hemiplegia or Paraplegia	Renal Disease	Malignancy, Except Skin Neoplasms	Moderate or Severe Liver Disease	Metastatic Solid Tumor	AIDS HIV
ICD Only	0.0%	1.0%	5.5%	3.5%	0.1%	12.0%	0.5%	0.0%	9.2%	0.0%	0.0%	1.9%	0.1%	0.0%	0.2%	0.8%	0.0%
SNOMED CT Only	0.0%	0.2%	2.3%	0.3%	1.4%	8.3%	0.1%	2.4%	0.1%	2.3%	5.5%	0.6%	10.9%	3.3%	0.7%	0.0%	0.2%
Neither	85.3%	73.4%	73.0%	74.9%	88.9%	38.5%	93.4%	90.7%	79.5%	61.3%	71.8%	91.1%	68.2%	84.8%	96.6%	96.2%	98.3%
Both	14.7%	25.4%	19.2%	21.3%	9.6%	41.2%	6.1%	6.9%	11.2%	36.4%	22.7%	6.4%	20.9%	11.9%	2.5%	3.0%	1.5%

Figure 2. Overlap in patient capture for each comorbidity comprising the CCI in DOD



	Myocardial Infraction	Congestive Heart Failure	Peripheral Vascular Disease	Cerebrovascular Disease	Dementia	Chronic Pulmonary Disease	Rheumatic Disease	Peptic Ulcer Disease	Mild Liver Disease	Diabetes Without Chronic Complication	Diabetes With Chronic Complication	Hemiplegia or Paraplegia	Renal Disease	Malignancy, Except Skin Neoplasms	Moderate or Severe Liver Disease	Metastatic Solid Tumor	AIDS HIV
ICD Only	0.0%	1.7%	13.9%	3.5%	0.1%	9.1%	1.0%	0.0%	11.0%	0.0%	0.0%	1.0%	0.0%	0.0%	0.1%	1.0%	0.0%
SNOMED CT Only	0.0%	0.3%	2.2%	0.3%	0.8%	10.9%	0.1%	2.8%	0.1%	1.3%	5.6%	0.4%	7.9%	8.7%	0.5%	0.0%	0.1%
Neither	82.7%	70.8%	63.8%	68.0%	87.9%	44.8%	91.6%	90.8%	82.4%	59.3%	70.3%	94.2%	62.7%	69.6%	97.7%	94.2%	99.5%
Both	17.3%	27.3%	20.2%	28.2%	11.2%	35.2%	7.3%	6.4%	6.4%	39.3%	24.1%	4.3%	29.3%	21.7%	1.7%	4.8%	0.4%

Table 1. Predictive performance of models to predict 1-year mortality

Model	MDCD, c statistic (95% CI)	DOD, c statistic (95% CI)
1	0.758 (0.755, 0.761)	0.753 (0.752, 0.754)
2	0.752 (0.749, 0.754)	0.757 (0.756, 0.758)
3	0.758 (0.755, 0.76)	0.76 (0.759, 0.761)
4	0.786 (0.784, 0.789)	0.802 (0.8, 0.803)
5	0.79 (0.787, 0.792)	0.804 (0.803, 0.805)

As indicated by the c-statistic, model 3 (SNOMED CT) slightly outperformed model 2 (ICD). While models 4 and 5 had similar performance, they significantly outperformed models 1, 2, and 3.

Conclusions

The current study found the SNOMED CT Disease Hierarchy was associated with a higher average CCI as compared to the Quan adaptation. Differences in patient capture were especially pronounced for the following comorbidities: chronic pulmonary disease, diabetes with chronic complications, renal disease, malignancy, peripheral vascular disease and mild liver disease. Nevertheless, both adaptations had similar performance in predicting 1-year mortality suggesting both versions represent comparable measures of clinical prognosis for risk adjustment.