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

Stephen P Fortin
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 Clininformatics 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 visit = first inpatient visit)

Covariates: The CCI and each comorbid condition comprising the CCI was measured based on all observed diagnosis codes recorded at or anytime 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 characteristic curve.

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

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

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

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.

<table>
<thead>
<tr>
<th></th>
<th>MDCD</th>
<th>DOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCI, SNOMED CT vs. ICD</td>
<td>3.91 vs. 4.12</td>
<td>4.43 vs. 4.6</td>
</tr>
</tbody>
</table>

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