Adaptation and Validation of the Charlson Comorbidity Index in Administrative Claims Data Using the SNOMED CT Standardized Vocabulary

Stephen P Fortin, Jenna Reps, Patrick Ryan

Background

The Charlson comorbidity index (CCI) provides a single aggregate measure of patient comorbidity and is commonplace in epidemiology research.[1] Multiple adaptations of the CCI have emerged for application to administrative claims data using the International Classification of Diseases, Ninth and Tenth Revisions (ICD-9/10) and their clinical modifications (ICD-9/10-CM).[2-5] While the OHDSI community has implemented a coding algorithm for the CCI using the SNOMED CT standardized vocabulary, no prior literature exists describing the development and validation of the OHDSI coding algorithm. Furthermore, previous studies have shown large discrepancies in the measurement of comorbidities between the OHDSI and Quan adaptations of the CCI.[6-7] To address the limitations of the OHDSI adaptation, the current study adapted and validated a new coding algorithm for the CCI using the SNOMED CT standardized vocabulary, henceforth referred to as the SNOMED adaptation.

Methods

The SNOMED adaptation was developed through the direct translation of the Quan coding algorithms followed by manual curation by clinical subject matter experts. Code mapping diagnostics were produced to identify the cause of all discrepant codes between coding algorithms. To validate the newly-developed algorithm, all inpatient visits occurring during the calendar years of 2013 and 2018 in two large U.S. administrative claims databases were identified; specifically, the Optum® De-Identified Clinformatics Data Mart Database – Date of Death (DOD) database; and IBM® MarketScan® Multi-State Medicaid Database (MDCD). Differences in the overall CCI and frequency of individual comorbid conditions comprising the CCI measured using the SNOMED versus Quan adaptations were assessed using standardized mean differences (SMD). Logistic regression was used to predict one-year mortality among hospitalized patients using the CCI as the dependent variable, and the performance of the SNOMED versus Quan adaptations were compared based on the c-statistic.

Results

A total of 5,343 ICD-9/10-CM codes mapping to either the SNOMED or Quan adaptations were identified, among which 4,646 (87.0%) codes were included in both algorithms and 695 (13.0%) represented discrepant codes. The mapping of multiple ICD codes to a single SNOMED CT code was the most common cause of discrepant codes (n=560; 80.6%), which resulted in the additional capture of clinically relevant codes in 24.6% (n=138) of cases. A higher prevalence of discrepant codes was observed among the following comorbidities comprising the CCI: rheumatic disease (n=130) and diabetes with chronic complications (n=211).

For each database-calendar year combination, as indicated by a SMD <0.10, no significant differences in the overall CCI were observed between the SNOMED versus Quan adaptations (MDCD, 2013: 3.75 versus 3.60; DOD, 2013: 3.63 versus 3.51; MDCD, 2018: 4.04 versus 3.91; DOD, 2018: 4.55 versus 4.43). Similarly,

no significant differences in the frequency individual comorbidities comprising the CCI was observed between vocabularies. The overall CCI and frequency of individual comorbidities among patients identified during the calendar year of 2018 for each database is shown in Table 1. As shown in Table 2, no significant differences in the performance of the SNOMED and Quan adaptations in predicting one-year mortality among hospitalized patients was observed.

	MDCD (N=491,311)			DOD (N=1,109,389)		
	SNOMED	**Quan**	**SMD**	**SNOMED**	**Quan**	**SMD**
Charlson Comorbidity Index, mean (sd)	4.04 (4)	3.91 (3.92)	0.023	4.55 (3.93)	4.43 (3.86)	0.022
Comorbid Conditions, n (%)						
Myocardial Infraction	72069 (14.7)	72057 (14.7)	0.000	192333 (17.3)	192333 (17.3)	0.000
Congestive Heart Failure	129372 (26.3)	129712 (26.4)	0.001	320600 (28.9)	321007 (28.9)	0.001
Peripheral Vascular Disease	123268 (25.1)	121219 (24.7)	0.007	383111 (34.5)	378005 (34.1)	0.007
Cerebrovascular Disease	121917 (24.8)	121824 (24.8)	0.000	351730 (31.7)	351659 (31.7)	0.000
Dementia	49312 (10)	47300 (9.6)	0.011	127958 (11.5)	125035 (11.3)	0.006
Chronic Pulmonary Disease	261239 (53.2)	261161 (53.2)	0.000	491536 (44.3)	491425 (44.3)	0.000
Rheumatic Disease	37012 (7.5)	32366 (6.6)	0.031	107692 (9.7)	92573 (8.3)	0.038
Peptic Ulcer Disease	33855 (6.9)	33855 (6.9)	0.000	70761 (6.4)	70761 (6.4)	0.000
Mild Liver Disease	103234 (21)	100175 (20.4)	0.012	198534 (17.9)	193778 (17.5)	0.008
Diabetes without Chronic Complication	180707 (36.8)	178911 (36.4)	0.005	441097 (39.8)	436760 (39.4)	0.005
Diabetes with Chronic Complication	123869 (25.2)	111701 (22.7)	0.044	293174 (26.4)	267237 (24.1)	0.039
Hemiplegia or Paraplegia	41778 (8.5)	40428 (8.2)	0.008	60555 (5.5)	58900 (5.3)	0.005
Renal Disease	106497 (21.7)	102761 (20.9)	0.014	330026 (29.8)	325401 (29.3)	0.006
Malignancy, Except Skin Neoplasms	60499 (12.3)	58426 (11.9)	0.010	247915 (22.4)	241187 (21.7)	0.011
Moderate or Severe Liver Disease	15374 (3.1)	13194 (2.7)	0.022	24232 (2.2)	19882 (1.8)	0.023
Metastatic Solid Tumor	19762 (4)	18820 (3.8)	0.008	66772 (6)	64521 (5.8)	0.007
AIDS/HIV	7471 (1.5)	7471 (1.5)	0.000	4850 (0.4)	4850 (0.4)	0.000

Table 1. Overall CCI and frequency of individual comorbidities among hospitalized patients in MDCD and DOD in 2018

Year	Adaptation of CCI	MDCD, c-statistic (95% CI)	DOD, c-statistic (95% CI)
2013	SNOMED	0.725 (0.721, 0.728)	0.789 (0.787, 0.79)
	Quan	0.723 (0.72, 0.726)	0.787 (0.786, 0.789)
2018	SNOMED	0.754 (0.751, 0.757)	0.757 (0.755, 0.758)
	Quan	0.752 (0.749, 0.754)	0.757 (0.756, 0.758)

Table 2. Performance of the SNOMED versus Quan adaptations of the CCI in predicting one-year mortality among hospitalized patients in MDCD and DOD in 2013 and 2018

Conclusion

The SNOMED adaptation had similar performance to the Quan adaptation in terms of measuring the overall CCI, frequency of individual comorbidities comprising the CCI, and predicting one-year mortality among hospitalized patients. Given the SNOMED adaptation permits for increased reproducibility and transparency of research, we posit the SNOMED adaptation as a substantial improvement to the current implementation of the CCI used by OHDSI.

References/Citations

1. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation. Journal of Chronic Diseases. 1987;40:373–83.
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. doi: 10.1016/0895-4356(92)90133-8. PMID: 1607900.
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. doi: 10.1016/0895-4356(93)90103-8. PMID: 8410092.
4. D'Hoore W, Bouckaert A, Tilquin C. Practical considerations on the use of the Charlson comorbidity index with administrative data bases. J Clin Epidemiol. 1996 Dec;49(12):1429-33. doi: 10.1016/s0895-4356(96)00271-5. PMID: 8991959.
5. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005 Nov;43(11):1130-9. doi: 10.1097/01.mlr.0000182534.19832.83. PMID: 16224307.
6. Fortin SP. Predictive performance of the Charlson Comorbidity Index: SNOMED CT disease hierarchy versus international classification of diseases. OHDSI. 2021.
7. Viernes B, Lynch KE, Robison B, Gatsby E, DuVall SL, Matheny ME. SNOMED CT disease hierarchies and the Charlson Comorbidity Index (CCI): an analysis of OHDSI methods for determining CCI. OHDSI. 2020.