A Comparison of the Strength of Associations Between Variants of the Charlson Comorbidity Index and Patient-Reported Health Outcomes

Edward A. Witt, PhD, Manager, Health Outcomes Research, formerly of Kantar Health, Princeton, NJ USA, and **Amir Goren, PhD**, Senior Director, Health Outcomes, Kantar Health, New York, NY, USA

Edward A. Witt, PhD

KEY POINTS . . .

Measures such as the Charlson Comorbidity Index (CCI) are often used in cross-sectional studies to control for non-relevant variance to the focal hypothesis of interest.

This analysis examined a few common versions of CCI scoring that can be applied readily to patient-reported data to assess their associations with patient-reported health outcomes.

All variants of the CCI examined had significant correlations with the outcomes, with variations in the strength of each, as well as caveats as to the context in which each might be used.

When looking at correlations between variables (e.g., the presence vs. absence of a disease) and outcomes (e.g., quality of life) in population-based cross-sectional studies, it is difficult to know how much of that correlation is attributable to unrelated influences (e.g., the fact that patients with the disease are also likely to have other diseases that impair their quality of life). Therefore, it is important to adjust in some fashion for these other potential influences.

The Charlson Comorbidity (CCI) Index is a weighted index of 19 comorbidities developed as a prognostic taxonomy to account for short-term risk of mortality from comorbid diseases in longitudinal studies [1]. The CCI, however, is increasingly being used in outcomes research studies as a means of controlling for comorbidities in multivariable regression models. In these studies, the outcome of interest is not necessarily mortality—instead, it may be economic or humanistic. Thus, the question of whether the CCI can account for meaningful variance in these measures is an important one.

Moreover, many alternative methods of scoring the CCI have been proposed relatively recently in the literature. In 2008, Charlson et al. added conditions to the original CCI in an effort to predict health care resource use, as opposed to clinical outcomes [2]. Independently, Quan et al. (2011) developed a scheme that removed or updated the weighting of comorbidities from the original CCI to reflect modern estimates of contributions to mortality risk in a broader sample of patients and hospitals [3]. Similar efforts have been undertaken to improve the CCI in clinical contexts [4].

Accordingly, the goal of this analysis was to evaluate the performance of the original CCI, as well as the two alternative scorings, in accounting for variance in humanistic patient-reported outcomes (i.e., quality of life, work productivity, and resource use, in a broad, real-world survey of US adults).

Study Details

Data for this analysis were extracted from the 2013 U.S. National Health and Wellness Survey (NHWS; n = 75,000). The NHWS is an annual, representative general health survey of US adults (≥ 18 years of age). CCI variants (identified by year as CCI, CCI 2008, and CCI 2011) were scored using self-reported diagnoses of conditions (rather than ICD-9 codes, which are not feasible using patient-reported data). Paraplegia and moderate/severe liver disease were not included in the scoring because precise questions about these conditions were not included in the NHWS. Health related quality of life (HRQoL) including mental and physical component summary scores (MCS and PCS, respectively), was assessed via the Short Form (SF)-36v2 [5], with health utilities derived using the SF-6D algorithm [6]. Productivity impairment (among employed respondents and in general) was assessed via the Work Productivity and Activity Impairment (WPAI-GH) questionnaire [7]; and health care resource use was assessed via self-reported counts in the last six months (i.e., number of emergency room (ER) visits, hospitalizations, and health care provider (HCP) visits). Pearson correlations (r) were used to examine variance explained in HRQoL, whereas Spearman correlations (ρ) were used to estimate variance explained in productivity impairment and health care resource use.

What We Found

Mathematically, CCI scores from all three variants can range from 0 to a high score of 26-42, depending on the number of comorbidities assessed (see Table 1). In the current study, most patients (95.4% to 98.8%) had a score between 0-3 across all three variants. CCI 2008 (incorporating the most comorbidities) yielded the score with the greatest variability, while CCI 2011 (incorporating the fewest comorbidities) had the least. The modal score in all three cases was 0.

CCI variants were all highly inter-correlated, with the highest correlation between CCI and CCI 2008 (r = 0.88) and the lowest between CCI 2008 and CCI 2011 (r = 0.75), with CCI and CCI 2011 in-between (r = 0.87).

>

Table 1. Comorbidities and weighting across CCI, CCI 2008, and CCI 2011.

Condition	CCI included (weight)	CCI 2008 included (weight)	CCI 2011 included (weight)
Chronic pulmonary disease	✓ (1)	✓ (1)	✓ (1)
Rheumatologic disease	✓ (1)	✓ (1)	✓ (1)
Diabetes without chronic complications	✓ (1)	✓ (1)	
Congestive heart failure	✓ (1)	✓ (1)	✓ (2)
Dementia	✓ (1)	✓ (1)	✓ (2)
Mild liver disease	✓ (1)	✓ (1)	✓ (2)
Myocardial infarction	✓ (1)	✓ (1)	
Peripheral vascular disease	✓ (1)	✓ (1)	
Cerebrovascular disease	✓ (1)	✓ (1)	
Peptic ulcer disease	✓ (1)	✓ (1)	
Diabetes with chronic complications	✓ (2)	✓ (2)	✓ (1)
Renal disease	✓ (2)	✓ (2)	✓ (1)
*Hemiplegia or paraplegia**	✓ (2)	✓ (2)	✓ (2)
Any tumor	✓ (2)	✓ (2)	✓ (2)
Leukemia	✓ (2)	✓ (2)	
Lymphoma	✓ (2)	✓ (2)	
*Moderate or severe liver disease**	✓ (3)	✓ (3)	✓ (4)
AIDS/HIV	✓ (6)	✓ (6)	✓ (4)
Metastatic solid tumor	✓ (6)	✓ (6)	✓ (6)
Hypertension		✓ (1)	
Depression		✓ (1)	
Use of warfarin		✓ (1)	
Skin ulcers/cellulitis		✓ (2)	
Total Conditions	19	23	14
Total Conditions in NHWS	17	21	12

**CCI was calculated on the basis of self-reported diagnosis with the various comorbidities of interest. Summary scores were created using the weights noted above. Excluded from the calculations were paraplegia and moderate/severe liver disease, as those conditions were not assessed in the NHWS.*

Correlations between CCI variants and health outcomes are presented in Figure 1.

CCI 2008 was the most strongly associated with all health outcomes examined (see middle column of Fig. 1), especially more so with multiple component summary (MCS), than its counterparts.

The original CCI (left column Fig. 1) was, on average, more strongly associated with outcomes than CCI 2011 (right column Fig. 1) in all but two cases—MCS and absenteeism—where differences were negligibly in favor of CCI 2011.

With regard to outcomes, all variants of CCI were associated most strongly with PCS, activity impairment, and HCP visits, whereas associations were much weaker for work productivity metrics, MCS, and extreme or rare events such as ER visits and hospitalizations.

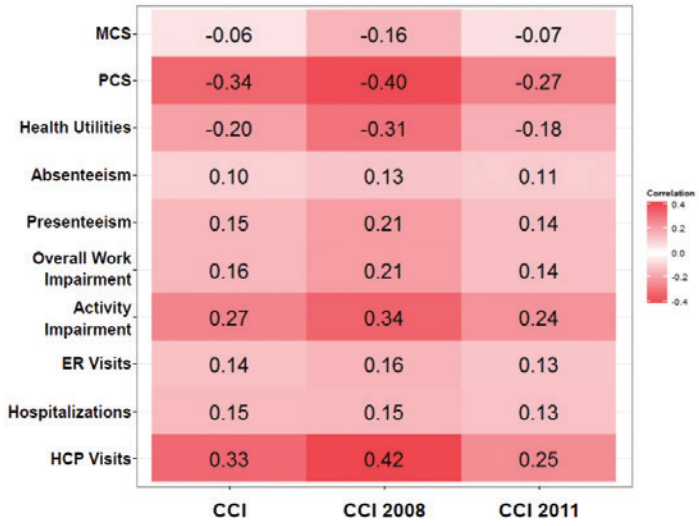
How Our Results Apply to Outcomes Studies

We examined three alternative scorings of the CCI in relation to three sets of humanistic health outcomes (health-related quality of life, work productivity and activity impairment, and health care resource use). All three scoring schemes resulted in strongly positively skewed distributions (median and modal scores of 0 for all three scorings); among these, CCI 2008 accounted for the most variance (i.e., the strongest associations with outcomes) across respondents.

CCI 2008 out-performed both the original CCI and CCI 2011 across health outcomes in terms of having the highest correlations—perhaps not surprising, given that it accounts for a larger number of comorbidities than either of its counterparts. Moreover, CCI 2008 includes depression, which undoubtedly contributed to its relatively strong correlation with MCS, an outcome measure assessing mental health status.

In general, CCI scores were better at accounting for variance in outcomes related to physical functioning (e.g., physical component summary [PCS], while accounting for less variance in work productivity (e.g., absenteeism, presenteeism) and extreme or rare events (e.g., ER visits, hospitalizations). Future CCI scoring schemes may be adapted to better account for variance in these other outcomes. For example, CCI 2008, to the extent that it accounts for pre-existing or chronic depression, helps control for variance in mental HRQoL that one may wish to control for when looking at the burden of an unrelated or physical condition. Future CCI schemes that account successfully for other pre-existing mental health conditions (e.g., anxiety disorders, attention deficit disorders) can help improve variance explained. A challenge with such an approach is ensuring that these assessments do not inadvertently control for mental health conditions that are consequences of the conditions of interest (e.g., depression arising as a consequence of another condition). This is easier to assess and accept in longitudinal studies, but becomes challenging in cross-sectional data. Similarly, CCI schemes that account for other conditions with implications for productivity impairments and more intense resource use (e.g., chronic migraines,

Figure 1. Heat map of CCI correlations with health outcomes.



ER indicates emergency room; HCP indicates health care provider; MCS indicates multiple component summary score; and PCS indicates physical component summary.

various types of ongoing pain) will help in accounting for additional, useful variance.

Considering Strengths and Limitations of Our Approach

This study used a large (n = 75,000) representative sample of US adults and to our knowledge is the first study to compare the relative validity of CCI variants in accounting for variance in patient-reported outcomes.

All variants of the CCI in the current study make use of a simple summary index of standard comorbidities rather than including all individual comorbidities simultaneously, and they may or may not capture all comorbidities of interest. This approach can be desirable in multivariable modeling when considerations of degrees of freedom come into play, as is the case where sample sizes are small and power is too limited to include all individual comorbidities as separate covariates. This added power, however, comes at the expense of maximizing variance explained.

One potential limitation of this work is that it used self-reported diagnoses rather than clinically verifiable ones (e.g., ICD-9 codes). The reliability of any index, however, is subject to the quality of the data used to compute it, and database records are not exempt from this limitation (e.g., incentives/motivation and accuracy of reporting can vary at both the time of original recording and later reporting/interpretation of the records).

A second limitation is that this work did not include paraplegia or moderate/severe liver disease in the calculation of the CCI, as the NHWS does not assess these comorbidities. The exclusion of paraplegia (because it is included in all variants and weighted the same) could only have affected absolute differences in CCI scores (i.e., means), not relative differences in variance explained across variants. The exclusion of moderate/severe liver disease, however, could have affected both absolute differences in CCI scores (i.e., means) and relative differences in variance explained (because it is assigned a greater weight in CCI 2011).

Finally, NHWS may under-represent elderly adults in higher age brackets (i.e., ≥75 years old) and those who have no access to the Internet and/or lack motivation to participate in panels and online surveys. Also, these results may not generalize to a more specialized population.

Conclusions and Recommendations

If researchers are interested in controlling for the most variance in these humanistic health outcomes via CCI, then CCI 2008 may be their chosen scheme. Variance explained, however, should not be the only decision rule applied to selection of covariates. For reasons noted above, CCI 2008 may be best adapted to longitudinal studies, administrative data-based studies (as opposed to patient-reported), and/or studies where current mental health status is of particular interest as a potential covariate.

For instance, although the study by Quan et al. [3] did not explain the most variance, it is based on more current data using broader samples than the other two variants, and it requires the fewest conditions to compute, which can help when the length of the survey is an important consideration.

Of course, this comes at the cost of reducing variance explained and risking having a more liberal interpretation of burden-of-illness effects when controlling for less variance. Therefore, the original CCI provides a more conservative and widely used, if somewhat outdated, version.

"If researchers are interested in controlling for the most variance in these humanistic health outcomes via CCI, then CCI 2008 may be their chosen scheme."

Finally, both practical and theoretical considerations should apply to selection or any modification of a particular variant of the CCI. For example, one should account for whether the condition of interest is among the CCI components, perhaps necessitating exclusion of that condition from the scoring. Otherwise, one should examine the relevance of comorbidities accounted for by the CCI to the condition or outcomes of interest, thus driving determination of which comorbidities are most appropriate to include.

References

[1] Charlson ME, Pompei P, Alex KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83. [2] Charlson ME, Charlson RE, Peterson JC, et al. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. J Clin Epidemiol 2008;61:1234-40. [3] Quan H, Li B, Couris CM, et al. Updating and validating the Charlson Comorbidity Index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. Am J Epidemiol 2011;173:676-82. [4] Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. J Clin Epidemiol 2000; 53:1258-67 [5] Maruish ME, editor. User's manual for the SF-36U2 Health Survey. Quality Metric Incorporated; 2011. [6] Brazier J, Roberts J, Deverill, M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271-92. [7] Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics 1993;4:353-65.