Navidence. {NESTED} KNOWLEDGE PurpleLab®

Poster Code: B-263

Scan for poster + more

Variation in Charlson Comorbidity Index Scores Using Different ICD-10 Code Adaptations: A Call for Harmonization in U.S. Real-World Data Research

Aimee Harrison MFA,[1] Jared Kamauu BFA,[1] Scott L. DuVall PhD,[2] Kevin Kallmes JD MA,[3] Allise Kamauu MS,[1] Garik Galustjan MS,[3] Michael Buck PhD,[1] Craig Parker MD MS,[1] Aaron Kamauu MD MS MPH[1]

[1]Navidence Inc., USA; [2]PurpleLab Healthcare Analytics, USA; [3]Nested Knowledge, USA

DISCLOSURES: AHarrison was an employee of Navidence during this work. JKamauu, AKamauu, MBuck, CParker, and AKamauu are employees of Navidence. SDuVall is an employee of PurpleLab. KKallmes and GGalustjan are employees of Nested Knowledge.

So What?

- *In U.S. data, ICD-10-WHO–based CCI adaptations can systematically under-represent clinical complexity, threatening study validity and comparability.*
- ***We call for harmonization to the Beyrer adaptation when calculating CCI in U.S. RWD.***
- *Researchers should explicitly report which CCI adaptation and specific code lists were used to ensure transparency and reproducibility.*

BACKGROUND

Context

The Charlson Comorbidity Index (CCI) is widely used in pharmacoepidemiologic research. As noted by Charlson (2022), *multiple adaptations* of ICD-10 codelist for CCI exist spanning international terminologies. (see

These adaptations were developed against different code systems — some for ICD-10-WHO and its national modifications, and one for ICD-10-CM as used in the United States.

*See companion poster **B-272***

The Gap / Problem

Despite fundamental differences between ICD-10-WHO and ICD-10-CM coding systems, researchers frequently apply these adaptations to U.S. real-world data (RWD) without appreciation for code system compatibility. In addition, many studies that include CCI do not publish the specific adaptation used; leaving it ambiguous to the audience.

Pharmacoepidemiologic Relevance

Mismatched code definitions can introduce systematic measurement error into a covariate used for confounding control in nearly every RWD study, affecting both study validity and comparability across studies. Because CCI is reported routinely in evidence submitted to regulators and payers, the choice of adaptation — and whether it is reported at all — has direct implications for the interpretability and reproducibility of pharmacoepidemiologic findings.

OBJECTIVES

Primary Objective

To quantify variation in CCI scores when applying different ICD-10 code adaptations to the same U.S. elderly patient cohorts.

Secondary Objectives

- Compare four published CCI adaptations — Beyrer 2021 (ICD-10-CM), Halfon 2002 (WHO), Sundararajan 2004 (WHO), and Quan 2005 (WHO) — applied with identical weighting.
- Quantify differences in the resulting CCI score distributions between adaptations.
- Assess the likelihood of comorbidity detection across three common thresholds (CCI 0 vs >0; <5 vs 5+; and <10 vs 10+).
- Assess score change magnitude for all CCI values >0.

LIMITATIONS

- Cross-sectional design: CCI scores were compared at a single time point and were not linked to downstream outcomes, so this analysis quantifies measurement differences rather than predictive performance.
- No reference standard for true comorbidity burden exists, so agreement between adaptations cannot by itself establish which adaptation is correct.
- Results reflect persons age 65+ in one U.S. claims database and one ICU-based EHR database, and may not generalize to younger or non-U.S. populations.
- Identical weights were applied to isolate the effect of code definitions; alternative weighting schemes proposed alongside some adaptations were not evaluated.
- Differences in coding intensity and care setting between claims and EHR sources may contribute to observed differences alongside the code definitions themselves.

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):373–383.
2. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson Comorbidity Index: a critical review of clinimetric properties. Psychother Psychosom. 2022;91(1):8–35.
3. Beyrer J, Manjelievskaia J, Bonafede M, et al. Validation of an International Classification of Disease, 10th revision coding adaptation for the Charlson Comorbidity Index in United States healthcare claims data. Pharmacoepidemiol Drug Saf. 2021;30(5):582–593.
4. Halfon P, Eggli Y, van Melle G, Chevalier J, Wasserfallen JB, Burnand B. Measuring potentially avoidable hospital readmissions. J Clin Epidemiol. 2002;55(6):573–587.
5. Sundararajan V, Henderson T, Perry C, Muggivan A, Quan H, Ghali WA. New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality. J Clin Epidemiol. 2004;57(12):1288–1294.
6. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130–1139.

METHODS

Study Design & Data Source

Cross-sectional study of persons age 65+ using PurpleLab® CLEAR Claims (N = 70,176,713) and 65+ w/ inpatient encounter using MIMIC IV EHR data (N = 21,542).

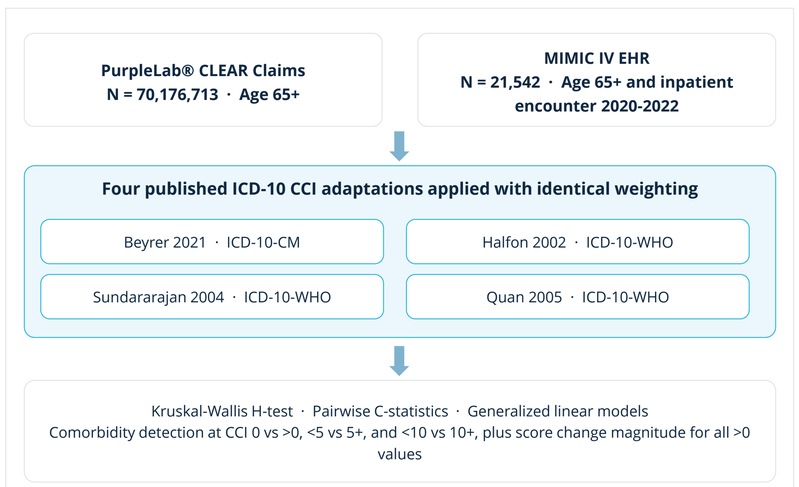
The two sources contrast a large national U.S. administrative claims population with a deeply coded, ICU-based EHR population.

The four published ICD-10 CCI adaptations had differences in the code lists used for defining comorbidities under CCI.

Study Population

All persons age 65 and older in each source were included: N = 70,176,713 in PurpleLab® CLEAR Claims and N = 21,542 in MIMIC IV (which also required ≥1 inpatient encounter between 2020-2022.

CCI was calculated for every patient under each of the four adaptations, so each patient serves as their own comparison across code definitions.



Study design. The same patients in each data source were scored under all four ICD-10 CCI adaptations using identical weights, isolating the effect of the code definitions themselves on the resulting CCI distribution.

Code Adaptations & Outcome Measures

Four published CCI adaptations were compared: Beyrer 2021 (ICD-10-CM), Halfon 2002 (WHO), Sundararajan 2004 (WHO), and Quan 2005 (WHO).

Outcomes: CCI score per patient using identical weighting; comorbidity detection at three common thresholds (CCI 0 vs >0; <5 vs 5+; and <10 vs 10+); and score change magnitude for all >0 values.

Statistical Analysis

Using identical weighting, we calculated CCI scores for all patients and quantified differences in the resulting distributions using the Kruskal-Wallis H-test.

Pairwise C-statistics were calculated to measure effect size. Generalized linear models were used to assess the likelihood of comorbidity detection across the three thresholds and score change magnitude for all >0 values.

Computable Operational Definitions (CODefs™)

Study elements — exposures, outcomes, and covariates — were defined using Navidence CODefs™: rigorous, data source-agnostic algorithms that ensure precision and reproducibility across RWR.

TRANSPARENCY & REPRODUCIBILITY

Code lists for all four CCI adaptations were taken from the published sources listed under References. Identical CCI weights were applied across adaptations so that differences reflect code definitions alone.

PurpleLab® CLEAR Claims were accessed under license; MIMIC IV is available to credentialed researchers via PhysioNet.

RESULTS

70,176,713 — Claims Patients Age 65+

21,542 — EHR Patients Age 65+

p < 0.0001 — Kruskal-Wallis Test Across Adaptations

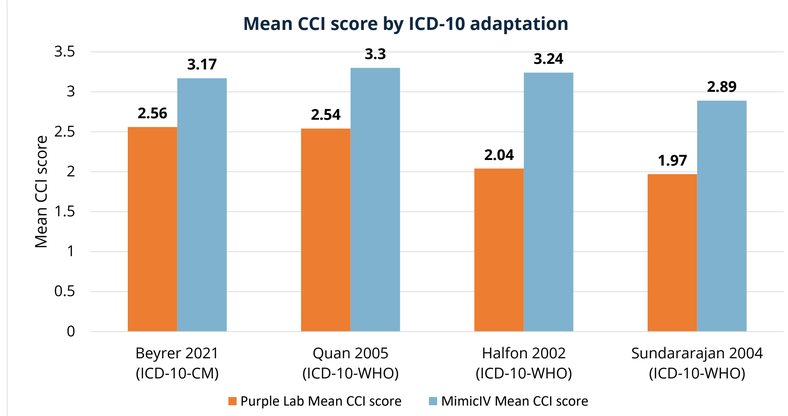


Figure 1. Mean CCI score under each of the four published ICD-10 adaptations, persons age 65+, PurpleLab® CLEAR Claims (N = 70,176,713) and inpatients age 65+ in MIMIC IV EHR (N=21,542). All adaptations were applied with identical weighting; differences across the four adaptations were significant (p < 0.0001).

CCI adaptation	Code system	Mean CCI (PL / MIMIC)	CCI 5+ (PL / MIMIC)	CCI 10+ (PL / MIMIC)
Beyrer 2021	ICD-10-CM	2.56 / 3.17	22.9% / 24.7%	4.0% / 4.3%
Quan 2005	ICD-10-WHO	2.54 / 3.30	22.6% / 27.7%	4.1% / 5.2%
Halfon 2002	ICD-10-WHO	2.04 / 3.24	15.6% / 27.1%	1.7% / 5.0%
Sundararajan 2004	ICD-10-WHO	1.97 / 2.89	15.8% / 23.9%	1.7% / 3.6%

Table 1. CCI results across the four published ICD-10 adaptations. Mean CCI, percentage of patients with CCI score of 5 or higher (CCI 5+) and of 10 or higher (CCI 10+) are reported from PurpleLab® CLEAR Claims (N = 70,176,713) and MIMIC IV (N = 21,542).

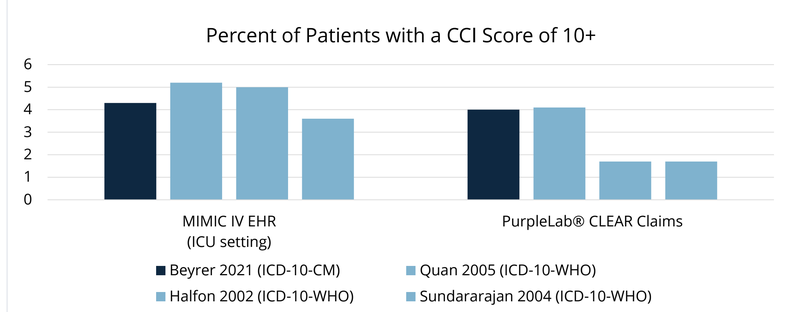


Figure 2. Percentage of patients with a CCI score of 10+ under the four CCI adaptations, in MIMIC IV EHR data (N = 21,542) and PurpleLab® CLEAR Claims (N = 70,176,713).

CONCLUSIONS

Key Findings

- Significant differences were found between the four CCI adaptations (p < 0.0001).
- In PurpleLab claims, Beyrer's ICD-10-CM adaptation captured the highest CCI scores (mean = 2.56), was closest to Quan (mean = 2.54, C-statistic = 0.498), and had statistically significantly (p < 0.001) greater odds of comorbidity detection from Sundararajan (mean = 1.97, C = 0.4552) and Halfon (mean = 2.04, C = 0.465).
- Similar results occurred in MIMIC data, except Beyrer produced higher percentages of patients with 10+ score than Halfon (27.7% vs 24.7%) than in PurpleLab claims (5.2% vs 4.3%) likely due to better capture of acute complications represented in ICD-10-CM and more completely captured in MIMIC's ICU setting.

Clinical & Policy Implications

Standardize on the Beyrer adaptation when calculating CCI in U.S. RWD, and report the adaptation and specific code lists used, so that comorbidity adjustment is transparent, comparable, and reproducible across studies.

Contact: aaron@navidence.com