Convergence of four measures of multi-morbidity

Brandon R. Grossardt, MS[1], Alanna M. Chamberlain, PhD[2,3], Cynthia M. Boyd, MD, MPH[4], William V. Bobo, MD, MPH[5], Jennifer L. St Sauver, PhD[2,6] and Walter A. Rocca, MD, MPH[2,7,8]

Abstract

Objectives: To compare the agreement between percentile ranks from 4 multi-morbidity scores.

Design: Population-based descriptive study.

Setting: Olmsted County, Minnesota (USA).

Participants: We used the medical records-linkage system of the Rochester Epidemiology Project (REP; http://www.rochesterproject.org) to identify all residents of Olmsted County, Minnesota who reached one or more birthdays between 1 January 2005 and 31 December 2014 (10 years).

Methods: For each person, we calculated 4 multi-morbidity scores using readily available diagnostic code lists from the US Department of Health and Human Services, the Clinical Classifications Software, and the Elixhauser Comorbidity Index. We calculated scores using diagnostic codes received in the 5 years before the index birthday and fit quantile regression models across age and separately by sex to transform unweighted, simple counts of conditions into percentile ranks as compared to peers of same age and of same sex. We compared the percentile ranks of the 4 multi-morbidity scores using intra-class correlation coefficients (ICCs).

Results: We assessed agreement in 181,553 persons who reached a total of 1,075,433 birthdays at ages 18 years through 85 years during the study period. In general, the percentile ranks of the 4 multi-morbidity scores exhibited high levels of agreement in 6 score-to-score pairwise comparisons. The agreement increased with older age for all pairwise comparisons, and ICCs were consistently greater than 0.65 at ages 50 years and older.

Conclusions: The assignment of percentile ranks may be a simple and intuitive way to assess the underlying trait of multi-morbidity across studies that use different measures.

[1]Division of Clinical Trials and Biostatistics, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA
[2]Division of Epidemiology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA
[3]Department of Cardiovascular Diseases, Mayo Clinic, Rochester, MN, USA
[4]Division of Geriatric Medicine and Gerontology, Department of Medicine, Johns Hopkins University, Baltimore, MD, USA
[5]Department of Psychiatry and Psychology, Mayo Clinic, Jacksonville, FL, USA
[6]The Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA
[7]Department of Neurology, Mayo Clinic, Rochester, MN, USA
[8]Women's Health Research Center, Mayo Clinic, Rochester, MN, USA

Corresponding author:
Walter A. Rocca, MD, MPH, Division of Epidemiology, Department of Quantitative Health Sciences, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.
Email: rocca@mayo.edu

Introduction

The accumulation of multiple chronic conditions (multimorbidity) as compared to peers of the same age and sex has been proposed as a clinical marker of acceleration of the aging processes.[1-4] We and others have previously shown that the number of chronic conditions increases steeply with older age, and may be influenced by factors such as sex, race, ethnicity, and socioeconomic status.[2,5-8] Thus, persons who have accumulated more chronic conditions than peers of the same age and sex may be experiencing accelerated aging.

If accumulation of chronic conditions is indeed a marker of accelerated aging, an agreed upon list of relevant chronic conditions across the lifespan would be optimal. Unfortunately, there is no consensus about how many conditions or clusters of conditions, or what severity of conditions, are needed to define multi-morbidity.[9] The variability in the numbers and types of conditions included in studies and the lack of age-, sex-, race-, and ethnicity-specific normative data hinders the comparison of study results from different research groups. We have recently reported normative data separately by age and sex in a well-defined US population using a multi-morbidity score of 20 chronic conditions defined by the US Department of Health and Human Services (DHHS).[10,11] Furthermore, we showed that the age- and sex-specific percentiles were predictive of 1-year and 5-year mortality. However, it is unclear how this measure compares with other measures of multi-morbidity.

To address the lack of consensus in multi-morbidity scoring methods, we calculated 4 multi-morbidity scores ranging from 18 conditions to 190 conditions in a well-defined US population.[10] After mapping the scores to age- and sex-specific percentiles, we investigated whether the 4 multi-morbidity scores had similar performance (agreement) in measuring a common underlying "construct" of multi-morbidity.[12,13]

Methods

Study population

We used the medical records-linkage system of the Rochester Epidemiology Project (REP) to identify all persons who lived in Olmsted County, Minnesota at any time between 1 January 2005 and 31 December 2014 (10 years). To be included, persons were required to have reached at least one birthday while a resident of the county within the study time frame; however, we excluded persons who did not provide authorization to use their medical record information for research, as per Minnesota legal requirements.[14] Persons in the sample were stratified into birthday cohorts at the single year level from age 0 (birth) through age 110 years, and regardless of the calendar year in which the birthday occurred. If persons resided in the county at several birthdays, they were included in the analyses multiple times (but always at different ages). Additional details about the study population were reported elsewhere.[10]

Scoring systems used to assess multi-morbidity

For each person, we calculated 4 multi-morbidity scores at each of their anchoring birthdays using readily available diagnostic code lists from the US Department of Health and Human Services (DHHS), the Clinical Classifications Software (CCS), and the Elixhauser Comorbidity Index. Table 1 provides a description of the 4 multi-morbidity scores included in our comparison study and the table in Supplemental Digital Content 1 provides a complete listing of the conditions.[10,15-22] All diagnostic codes were from the International Classification of Diseases, Ninth Revision (ICD-9).[23] Each birthday was treated as an index date, and prevalent conditions were determined from the diagnostic codes received by each person in the 5 years before the index date. For persons to qualify as having a specific prevalent condition, they were required to have 2 or more diagnosis codes separated by more than 30 days from among the list of condition-defining codes (see Table 1). For each of the 4 multi-morbidity scoring systems, a raw score was calculated as the simple unweighted count of the number of conditions present on the index date.

Statistical analyses

Quantile regression was used to model the raw multi-morbidity scores across age and separately within women and men. Because the multi-morbidity scores are integer values (counts) ranging between 0 and a maximum number of conditions (e.g., 20 or 190), we used methods for calculating quantiles for counts as described elsewhere.[24] The modeling involves a step of adding a random amount of jitter from a uniform distribution between 0 and 1 to each observed multi-morbidity score. We performed 30 replicates of the jitter process and averaged across the modeled

Table 1. The 4 multi-morbidity scores included in the comparison study.

Abbreviation Full score name	Number of conditions*	Description
DHHS US Department of Health and Human Services	20	The 20 conditions were selected by the DHHS because they are chronic, prevalent, and "potentially amenable to public health or clinical interventions or both."[10,15] However, we modified the set of ICD codes used to define cancer by excluding non-melanoma skin cancer. This is the multimorbidity score that we use in our previous normative study,[10] and in several previous publications on multi-morbidity in the REP.[2,5,6,9]
CCS Clinical Classifications Software chronic conditions	190	We constructed a multi-morbidity score based on the Clinical Classifications Software (CCS) enumeration of medical conditions.[16-17] The CCS groups each of the 15,072 billable ICD-9 diagnosis codes into one of 283 categories (i.e., conditions). From among all ICD-9 diagnosis codes, we identified 4,584 codes flagged as "chronic" via the Chronic Condition Indicator (CCI) tool.[18] We used the subset of chronic codes to identify the 190 chronic condition categories included in our multi-morbidity score. The 93 CCS categories with no diagnosis codes flagged as "chronic" were not included in the calculation of the CCS score.
CHAP Chapters of the CCS chronic conditions	18	We constructed a condensed version of the CCS multi-morbidity score by using only the 18 broad chapters into which all CCS categories are grouped. To define the chapters multi-morbidity score, the list of 4,584 ICD-9 diagnosis codes flagged as chronic were used from the CCS score.[16-18] The 18 CCS chapters roughly correspond to the body's key organ systems.
ELIX Elixhauser Comorbidity Index	31	We used the list of 31 conditions and the coding as described by Quan, et al. to define the Elixhauser Comorbidity Index.[19-22] The Elixhauser Comorbidity Index is one of the most commonly used and most studied measures of comorbidity.[19-22]

DHHS = Department of Health and Human Services, 20 condition score; ICD = International Classification of Diseases; REP = Rochester Epidemiology Project; CCS = Clinical Classifications Software, 190 chronic condition score; CHAP = Body system chapters (collapsed CCS system), 18 chapters score; ELIX = Elixhauser Comorbidity Index, 31 condition score.

*Many conditions are similar across the 4 different scores; however, the level of granularity differs among the scores. For example, hypertension is included in the DHHS list as one condition, whereas ELIX splits hypertension into two conditions (uncomplicated and complicated hypertension). The same is true for how ELIX splits diabetes into complicated and uncomplicated variants. The level of granularity varies greatly for cancer: the CCS score includes 33 separate types of cancer; the DHHS and CHAP scores group together all types of cancer into one category; the ELIX score includes 3 separate types of cancer for lymphoma, metastatic cancer, and solid tumors without metastasis. For more details, see Supplemental Digital Content 1.

quantiles to generate the final quantile profile for each score. Quantile regression was modelled using the *gcrq* function (growth charts regression quantiles) available as part of the *quantregGrowth* package in R. Although our agreement analyses focused on the quantiles in the adult population at ages 18 years through age 85 years, quantile regression models included birthdays from age 0 (birth) through age 95 to stabilize the younger (left side) and older (right side) ends of quantile regression curves.

The raw multi-morbidity scores of each person on a given birthday were transformed into percentile (quantile) ranks as compared to peers of same age and sex in the general population. Thus, the raw multi-morbidity scores derived from 4 lists with broadly different numbers of conditions were transformed into a common range from 1st percentile through 99th percentile. For simplicity, and because we found no significant differences in our previous study, we did not fit separate percentile regression models by race (Whites vs. non-Whites) or by ethnicity (Hispanics vs. non-Hispanics).[10] However, we did assess the agreement of the 4 scores separately in race and ethnicity strata.

We calculated intraclass correlation coefficients (ICC) assuming two-way random effects (i.e., a random set of included persons and a random set of multi-morbidity scores from amongst the universe of possibilities). In particular, we calculated the absolute agreement ICC using the *icc* function available as part of the *irr* (inter-rater reliability) package in R.[25] This ICC is given the notation of ICC(2,1) by Shrout and Fleiss and as ICC(A,1) by McGraw and Wong and in the *irr* package documentation.[12,13,25] The estimates of ICCs were accompanied by 95% confidence intervals. Scatter-plots of the pairwise multi-morbidity percentile scores were created (quantile-quantile plots) and included a reference line drawn on the one-to-one diagonal and a locally weighted regression (lowess) smoothed line. The lowess line provided a visual reference of deviation from the one-to-one diagonal line.[26]

To give a better visual representation of data density, a random amount of noise (jitter) in the range negative one to one (-1 to 1) was added to each percentile rank for quantile-quantile scatter plots. This was necessary because percentiles are predicted at integer values and many points would otherwise overlap. We conducted ICC analyses for women and men separately and combined. However, for simplicity, scatter plots are shown for women and men combined. All analyses were performed using SAS version 9.4 (SAS Institute) and R version 3.6.2.

Results

Study population

We identified 262,064 persons who resided in Olmsted County at some time between 1 January 2005 and 31 December 2014. Among those persons, 238,010 had reached at least 1 birthday while a resident and had given permission for the use of their medical records for research (see Figure, Supplemental Digital Content 2). We included persons more than once if they reached multiple birthdays during the 10-year period, but always at different ages. The total number of birthdays reached was 1,458,094, and the median number of birthdays was 6 per person (interquartile range of 3-10). However, we included in the quantile regression models only the 1,456,052 birthdays at age 0 (birth) through age 95 (among 237,791 unique persons). This age restriction ensured that quantile regression models were firmly anchored at younger ages and at older ages. The final agreement analyses were undertaken on the percentile ranks from 1,075,433 birthdays at ages 18 through 85 years (among 181,553 unique persons).

Percentile ranks of multi-morbidity

Table 2 shows the distribution by sex, race, ethnicity, and multi-morbidity scores in each of the birthday age cohorts from 18 to 85 years. The range of raw multi-morbidity scores varied substantially, particularly at older ages. Approximately half of the birthdays were in women (52.9%), and non-Whites and Hispanics were more common at younger ages than at older ages. These patterns are consistent with the underlying population of Olmsted County during the timeframe of the study.[27] At each birthday, the raw multi-morbidity scores were transformed into age- and sex-specific percentiles using quantile regression (see Figure, Supplemental Digital Content 3).

Table 3 shows the ICC values for 6 pairwise comparisons between the 4 multi-morbidity percentile ranks scores by age (in 5-year increments), for both sexes combined, and separately in strata for women and men. Overall, we found strong agreement of the 4 multi-morbidity scores with overall ICCs larger than 0.70 (Table 3). Figure 1 shows the quantile-quantile scatterplots for the 6 pairwise comparisons between the 4 multi-morbidity scores (both sexes pooled and for all birthdays at ages 18 to 85 years). The quantile-quantile scatter plots included a reference line on the one-to-one diagonal (black dashed line) and a reference line based on lowess smoothing (green solid line). The agreement varied across age and was higher at older ages (Table 3 and Figure 2). The lower ICC agreement values at younger ages are due to the narrow range of raw scores on scales with limited numbers of conditions. This narrow range of raw scores leads to large jumps in percentile ranks with the addition of a single condition. Therefore, the age effect was particularly strong for the pairwise comparison involving the DHHS and the CHAP scores (Figure 2). We also show quantile-quantile plots across three broad age strata to illustrate the impact of age on ICC values (see Figure, Supplemental Digital Content 4).

The table in Supplemental Digital Content 5 shows the ICC values for 6 pairwise comparisons between the 4 multi-morbidity percentile ranks scores by age (in 10-year increments), by race (non-White and White), and by ethnicity (Hispanic and non-Hispanic). The results from the 6 pairwise comparisons were similar across race and ethnic groups.

Discussion

Principal findings

After transformation of multi-morbidity scores into age- and sex-specific percentile ranks, the 4 scores exhibited high levels of agreement and supported the idea of a common underlying "multi-morbidity" construct. In particular, the 2 commonly used and clinically derived multi-morbidity scores, DHHS and Elixhauser, exhibited substantial agreement across all adult ages 18 through 85 years. As expected, we observed almost complete agreement between the 190 CCS chronic conditions score and the condensed 18 CCS Chapters (CHAP) score at all ages. The agreement between percentile ranks was lower at ages younger than 50 years when both scores included a small number of conditions (DHHS vs. CHAP), or when at least 1 score included a small number of conditions (DHHS vs. CCS). However, in general, the pairwise agreement across the 4 multi-morbidity percentile ranks was higher at ages 50 years and older.

We included scores ranging from 18 chapters of the CCS to 190 CCS chronic conditions. Our findings suggest that the percentile rank of a person compared to peers of same age and sex is a consistent measurement across 4 multi-morbidity scores, particularly at ages 50 years and older. The lack of agreement of multi-morbidity percentile ranks in younger

Table 2. Characteristics of persons included in each of the adult birthday age cohorts (ages 18 to 85 years).*

Age	Number of births	Female sex, N (%)	Non-white race, N (%)	Hispanic ethnicity, N (%)	Raw scores, (minimum, Q1, median, Q3, maximum)			
					DHHS	CCS	CHAP	ELIX
18	19,711	9,859 (50.0)	4,158 (21.1)	1,284 (6.5)	(0, 0, 0, 0, 5)	(0, 0, 0, 1, 17)	(0, 0, 0, 1, 9)	(0, 0, 0, 0, 7)
19	20,454	10,468 (51.2)	4,318 (21.1)	1,344 (6.6)	(0, 0, 0, 0, 5)	(0, 0, 0, 1, 18)	(0, 0, 0, 1, 9)	(0, 0, 0, 0, 7)
20	20,468	10,711 (52.3)	4,335 (21.2)	1,371 (6.7)	(0, 0, 0, 0, 6)	(0, 0, 0, 1, 23)	(0, 0, 0, 1, 10)	(0, 0, 0, 1, 11)
21	20,601	11,002 (53.4)	4,399 (21.4)	1,427 (6.9)	(0, 0, 0, 0, 6)	(0, 0, 0, 1, 29)	(0, 0, 0, 1, 10)	(0, 0, 0, 0, 12)
22	21,061	11,522 (54.7)	4,493 (21.3)	1,428 (6.8)	(0, 0, 0, 0, 7)	(0, 0, 0, 1, 32)	(0, 0, 0, 1, 12)	(0, 0, 0, 0, 13)
23	22,161	12,397 (55.9)	4,736 (21.4)	1,536 (6.9)	(0, 0, 0, 0, 7)	(0, 0, 0, 1, 34)	(0, 0, 0, 1, 14)	(0, 0, 0, 0, 14)
24	22,410	12,566 (56.1)	4,824 (21.5)	1,628 (7.3)	(0, 0, 0, 0, 7)	(0, 0, 0, 1, 25)	(0, 0, 0, 1, 13)	(0, 0, 0, 0, 9)
25	22,376	12,572 (56.2)	4,889 (21.8)	1,629 (7.3)	(0, 0, 0, 0, 8)	(0, 0, 0, 1, 24)	(0, 0, 0, 1, 13)	(0, 0, 0, 0, 10)
26	22,637	12,588 (55.6)	5,045 (22.3)	1,681 (7.4)	(0, 0, 0, 0, 9)	(0, 0, 0, 1, 25)	(0, 0, 0, 1, 14)	(0, 0, 0, 0, 12)
27	22,771	12,496 (54.9)	5,111 (22.4)	1,700 (7.5)	(0, 0, 0, 0, 10)	(0, 0, 0, 1, 26)	(0, 0, 0, 1, 13)	(0, 0, 0, 0, 12)
28	23,043	12,526 (54.4)	5,181 (22.5)	1,694 (7.4)	(0, 0, 0, 0, 10)	(0, 0, 0, 1, 26)	(0, 0, 0, 1, 13)	(0, 0, 0, 0, 11)
29	22,861	12,348 (54.0)	5,121 (22.4)	1,683 (7.4)	(0, 0, 0, 0, 10)	(0, 0, 0, 1, 32)	(0, 0, 0, 1, 12)	(0, 0, 0, 0, 12)
30	22,769	12,097 (53.1)	5,176 (22.7)	1,687 (7.4)	(0, 0, 0, 0, 10)	(0, 0, 0, 1, 41)	(0, 0, 0, 1, 14)	(0, 0, 0, 1, 14)
31	22,491	11,856 (52.7)	5,161 (22.9)	1,642 (7.3)	(0, 0, 0, 0, 8)	(0, 0, 0, 1, 27)	(0, 0, 0, 1, 12)	(0, 0, 0, 1, 14)
32	21,907	11,453 (52.3)	5,096 (23.3)	1,571 (7.2)	(0, 0, 0, 1, 9)	(0, 0, 0, 2, 32)	(0, 0, 0, 1, 12)	(0, 0, 0, 1, 15)
33	20,877	10,823 (51.8)	4,909 (23.5)	1,499 (7.2)	(0, 0, 0, 1, 9)	(0, 0, 0, 2, 36)	(0, 0, 0, 1, 13)	(0, 0, 0, 1, 15)
34	20,102	10,319 (51.3)	4,684 (23.3)	1,425 (7.1)	(0, 0, 0, 1, 9)	(0, 0, 0, 2, 29)	(0, 0, 0, 2, 10)	(0, 0, 0, 1, 15)
35	19,534	9,988 (51.1)	4,435 (22.7)	1,317 (6.7)	(0, 0, 0, 1, 8)	(0, 0, 0, 2, 28)	(0, 0, 0, 2, 12)	(0, 0, 0, 1, 12)
36	18,852	9,616 (51.0)	4,166 (22.1)	1,240 (6.6)	(0, 0, 0, 1, 9)	(0, 0, 0, 2, 30)	(0, 0, 1, 2, 12)	(0, 0, 0, 1, 13)
37	18,401	9,323 (50.7)	4,006 (21.8)	1,215 (6.6)	(0, 0, 0, 1, 10)	(0, 0, 1, 2, 34)	(0, 0, 1, 2, 13)	(0, 0, 0, 1, 14)
38	18,015	9,146 (50.8)	3,912 (21.7)	1,190 (6.6)	(0, 0, 0, 1, 9)	(0, 0, 1, 2, 34)	(0, 0, 1, 2, 14)	(0, 0, 0, 1, 14)
39	17,771	9,077 (51.1)	3,768 (21.2)	1,159 (6.5)	(0, 0, 0, 1, 9)	(0, 0, 1, 2, 29)	(0, 0, 1, 2, 14)	(0, 0, 0, 1, 14)
40	17,733	9,076 (51.2)	3,590 (20.2)	1,128 (6.4)	(0, 0, 0, 1, 10)	(0, 0, 1, 2, 28)	(0, 0, 1, 2, 12)	(0, 0, 0, 1, 14)
41	17,931	9,227 (51.5)	3,400 (19.0)	1,101 (6.1)	(0, 0, 0, 1, 10)	(0, 0, 1, 2, 29)	(0, 0, 1, 2, 12)	(0, 0, 0, 1, 14)
42	18,227	9,381 (51.5)	3,249 (17.8)	1,098 (6.0)	(0, 0, 0, 1, 11)	(0, 0, 1, 2, 30)	(0, 0, 1, 2, 11)	(0, 0, 0, 1, 15)
43	18,714	9,612 (51.4)	3,113 (16.6)	1,092 (5.8)	(0, 0, 0, 1, 12)	(0, 0, 1, 2, 32)	(0, 0, 1, 2, 13)	(0, 0, 0, 1, 16)
44	19,129	9,908 (51.8)	3,028 (15.8)	1,040 (5.4)	(0, 0, 0, 1, 12)	(0, 0, 1, 3, 34)	(0, 0, 1, 2, 14)	(0, 0, 0, 1, 17)
45	19,366	10,015 (51.7)	2,821 (14.6)	1,025 (5.3)	(0, 0, 0, 1, 12)	(0, 0, 1, 3, 37)	(0, 0, 1, 2, 14)	(0, 0, 0, 1, 16)
46	19,816	10,288 (51.9)	2,675 (13.5)	1,018 (5.1)	(0, 0, 0, 1, 12)	(0, 0, 1, 3, 36)	(0, 0, 1, 2, 13)	(0, 0, 0, 1, 15)
47	20,312	10,596 (52.2)	2,612 (12.9)	1,002 (4.9)	(0, 0, 0, 1, 12)	(0, 0, 1, 3, 37)	(0, 0, 1, 2, 13)	(0, 0, 0, 1, 16)
48	20,704	10,773 (52.0)	2,528 (12.2)	969 (4.7)	(0, 0, 0, 1, 12)	(0, 0, 1, 3, 38)	(0, 0, 1, 3, 13)	(0, 0, 0, 1, 16)
49	21,044	10,967 (52.1)	2,428 (11.5)	950 (4.5)	(0, 0, 0, 1, 12)	(0, 0, 1, 3, 35)	(0, 0, 1, 3, 13)	(0, 0, 0, 1, 13)
50	21,151	11,034 (52.2)	2,333 (11.0)	915 (4.3)	(0, 0, 1, 2, 11)	(0, 0, 1, 3, 33)	(0, 0, 1, 3, 12)	(0, 0, 0, 1, 14)
51	21,086	11,016 (52.2)	2,243 (10.6)	871 (4.1)	(0, 0, 1, 2, 12)	(0, 0, 1, 3, 33)	(0, 0, 1, 3, 13)	(0, 0, 0, 1, 16)
52	20,854	10,899 (52.3)	2,179 (10.4)	855 (4.1)	(0, 0, 1, 2, 13)	(0, 0, 2, 4, 37)	(0, 0, 2, 3, 13)	(0, 0, 0, 1, 19)
53	20,526	10,708 (52.2)	2,082 (10.1)	800 (3.9)	(0, 0, 1, 2, 13)	(0, 0, 2, 4, 39)	(0, 0, 2, 3, 13)	(0, 0, 0, 1, 18)
54	20,010	10,426 (52.1)	2,019 (10.1)	785 (3.9)	(0, 0, 1, 2, 13)	(0, 0, 2, 4, 38)	(0, 0, 2, 3, 14)	(0, 0, 1, 2, 16)
55	19,272	10,065 (52.2)	1,943 (10.1)	756 (3.9)	(0, 0, 1, 2, 12)	(0, 0, 2, 4, 36)	(0, 0, 2, 3, 13)	(0, 0, 1, 2, 17)
56	18,445	9,647 (52.3)	1,871 (10.1)	734 (4.0)	(0, 0, 1, 2, 12)	(0, 0, 2, 4, 40)	(0, 1, 2, 3, 13)	(0, 0, 1, 2, 17)
57	17,558	9,198 (52.4)	1,711 (9.7)	697 (4.0)	(0, 0, 1, 2, 14)	(0, 1, 2, 4, 43)	(0, 1, 2, 4, 14)	(0, 0, 1, 2, 18)
58	16,731	8,745 (52.3)	1,617 (9.7)	623 (3.7)	(0, 0, 1, 2, 14)	(0, 1, 2, 5, 43)	(0, 1, 2, 4, 14)	(0, 0, 1, 2, 18)
59	16,006	8,349 (52.2)	1,537 (9.6)	589 (3.7)	(0, 0, 1, 3, 13)	(0, 1, 2, 5, 47)	(0, 1, 2, 4, 15)	(0, 0, 1, 2, 18)
60	15,218	7,943 (52.2)	1,440 (9.5)	562 (3.7)	(0, 0, 1, 3, 13)	(0, 1, 3, 5, 49)	(0, 1, 2, 4, 14)	(0, 0, 1, 2, 18)
61	14,456	7,592 (52.5)	1,362 (9.4)	540 (3.7)	(0, 0, 1, 3, 12)	(0, 1, 3, 5, 34)	(0, 1, 2, 4, 14)	(0, 0, 1, 2, 20)
62	13,729	7,241 (52.7)	1,260 (9.2)	476 (3.5)	(0, 0, 2, 3, 13)	(0, 1, 3, 5, 48)	(0, 1, 3, 4, 14)	(0, 0, 1, 2, 20)
63	13,046	6,945 (53.2)	1,193 (9.1)	465 (3.6)	(0, 0, 2, 3, 13)	(0, 1, 3, 5, 41)	(0, 1, 3, 4, 13)	(0, 0, 1, 2, 17)
64	12,395	6,588 (53.2)	1,116 (9.0)	427 (3.4)	(0, 1, 2, 3, 13)	(0, 1, 3, 6, 43)	(0, 1, 3, 4, 13)	(0, 0, 1, 2, 18)
65	11,870	6,286 (53.0)	1,052 (8.9)	402 (3.4)	(0, 1, 2, 3, 12)	(0, 1, 3, 6, 44)	(0, 1, 3, 4, 13)	(0, 0, 1, 2, 18)
66	11,353	6,033 (53.1)	979 (8.6)	387 (3.4)	(0, 1, 2, 3, 12)	(0, 1, 4, 6, 43)	(0, 1, 3, 4, 13)	(0, 0, 1, 2, 18)
67	10,729	5,685 (53.0)	896 (8.4)	328 (3.1)	(0, 1, 2, 3, 13)	(0, 2, 4, 6, 48)	(0, 2, 3, 5, 13)	(0, 0, 1, 2, 18)
68	10,215	5,435 (53.2)	835 (8.2)	306 (3.0)	(0, 1, 2, 4, 13)	(0, 2, 4, 7, 39)	(0, 2, 3, 5, 13)	(0, 0, 1, 3, 17)

(continued)

Table 2. (continued)

Age	Number of birthdays	Female sex, N (%)	Non-white race, N (%)	Hispanic ethnicity, N (%)	Raw scores, (minimum, Q1, median, Q3, maximum)			
					DHHS	CCS	CHAP	ELIX
69	9,734	5,162 (53.0)	791 (8.1)	274 (2.8)	(0, 1, 2, 4, 12)	(0, 2, 4, 7, 38)	(0, 2, 3, 5, 14)	(0, 0, 1, 3, 16)
70	9,420	4,981 (52.9)	752 (8.0)	248 (2.6)	(0, 1, 3, 4, 13)	(0, 2, 4, 7, 39)	(0, 2, 3, 5, 14)	(0, 0, 1, 3, 18)
71	8,920	4,698 (52.7)	712 (8.0)	218 (2.4)	(0, 1, 3, 4, 13)	(0, 2, 5, 7, 44)	(0, 2, 4, 5, 13)	(0, 1, 2, 3, 18)
72	8,467	4,500 (53.1)	654 (7.7)	204 (2.4)	(0, 1, 3, 4, 13)	(0, 3, 5, 8, 33)	(0, 2, 4, 5, 13)	(0, 1, 2, 3, 19)
73	7,975	4,272 (53.6)	558 (7.0)	179 (2.2)	(0, 1, 3, 4, 13)	(0, 3, 5, 8, 34)	(0, 2, 4, 5, 13)	(0, 1, 2, 3, 17)
74	7,682	4,133 (53.8)	523 (6.8)	163 (2.1)	(0, 2, 3, 4, 14)	(0, 3, 5, 8, 34)	(0, 3, 4, 5, 13)	(0, 1, 2, 3, 17)
75	7,356	3,979 (54.1)	482 (6.6)	157 (2.1)	(0, 2, 3, 4, 15)	(0, 3, 6, 9, 38)	(0, 3, 4, 5, 13)	(0, 1, 2, 3, 17)
76	7,025	3,841 (54.7)	443 (6.3)	138 (2.0)	(0, 2, 3, 5, 15)	(0, 3, 6, 9, 40)	(0, 3, 4, 6, 12)	(0, 1, 2, 3, 19)
77	6,773	3,717 (54.9)	420 (6.2)	120 (1.8)	(0, 2, 3, 5, 14)	(0, 3, 6, 9, 37)	(0, 3, 4, 6, 12)	(0, 1, 2, 4, 18)
78	6,464	3,538 (54.7)	388 (6.0)	102 (1.6)	(0, 2, 3, 5, 13)	(0, 4, 6, 10, 40)	(0, 3, 4, 6, 14)	(0, 1, 2, 4, 19)
79	6,164	3,434 (55.7)	349 (5.7)	103 (1.7)	(0, 2, 3, 5, 14)	(0, 4, 6, 10, 32)	(0, 3, 4, 6, 12)	(0, 1, 2, 4, 17)
80	5,876	3,298 (56.1)	321 (5.5)	95 (1.6)	(0, 2, 4, 5, 13)	(0, 4, 7, 10, 36)	(0, 3, 4, 6, 12)	(0, 1, 2, 4, 16)
81	5,559	3,177 (57.2)	285 (5.1)	87 (1.6)	(0, 2, 4, 5, 15)	(0, 4, 7, 10, 38)	(0, 3, 5, 6, 13)	(0, 1, 2, 4, 18)
82	5,223	3,055 (58.5)	238 (4.6)	77 (1.5)	(0, 2, 4, 5, 14)	(0, 4, 7, 11, 36)	(0, 3, 5, 6, 12)	(0, 1, 3, 4, 17)
83	4,992	2,964 (59.4)	216 (4.3)	71 (1.4)	(0, 2, 4, 5, 15)	(0, 4, 7, 11, 36)	(0, 3, 5, 6, 12)	(0, 1, 3, 4, 17)
84	4,623	2,767 (59.9)	195 (4.2)	60 (1.3)	(0, 2, 4, 6, 14)	(0, 5, 7, 11, 34)	(0, 3, 5, 6, 13)	(0, 1, 3, 5, 16)
85	4,281	2,611 (61.0)	158 (3.7)	54 (1.3)	(0, 2, 4, 6, 15)	(0, 5, 8, 11, 39)	(0, 3, 5, 6, 14)	(0, 1, 3, 5, 20)
Ages 18-85	1,075,433	568,558 (52.9)	168,520 (15.7)	56,641 (5.3)	(0, 0, 0, 2, 15)	(0, 0, 1, 3, 49)	(0, 0, 1, 3, 15)	(0, 0, 0, 1, 20)

DHHS = Department of Health and Human Services, 20 condition score; CCS = Clinical Classifications Software, 190 chronic condition score; CHAP = Body system chapters (collapsed CCS system), 18 chapters score; ELIX = Elixhauser comorbidity index, 31 condition score; Q1 = 25th percentile (first quartile); Q3 = 75th percentile (third quartile).
*The characteristics of persons who experienced multiple birthdays during the study period were included multiple times (at different ages).

persons is somewhat expected because younger persons are generally healthier, have fewer medical visits, and have a limited range of observed raw multi-morbidity scores. In addition, most multi-morbidity scores include conditions common in the aging population (e.g., hypertension) but less frequently include conditions common in young persons (e.g., autism). For example, the DHHS score does not include obesity or anemia, whereas these conditions are included in the CCS and Elixhauser scores (and may both be common and important in younger persons). Thus, when transforming raw multi-morbidity scores to a percentile rank, the addition of 1 additional chronic condition leads to a marked jump in a person's percentile rank. In other words, multi-morbidity scores consisting largely of age-related conditions can only assess the upper tail of percentile ranks at younger ages. By contrast, persons at older ages have a broader range of observed raw multi-morbidity scores, allowing for more granularity in classifying persons into percentile ranks as compared to their peers.

Our analyses demonstrate strong agreement of the multi-morbidity percentile ranks for the 20 DHHS conditions as compared to the percentile ranks obtained using the more extensive list of 190 CCS chronic conditions and the 31 Elixhauser conditions. The agreement was particularly strong at ages 50 years and older. In summary, our study suggests that the 20 DHHS conditions are adequate in number and type to assign percentile ranks relative to peers of the same age and sex and to capture the underlying trait of multi-morbidity as well as the 190 CCS chronic conditions and the 31 Elixhauser conditions.

Unfortunately, there is no consensus on how many conditions, what clusters of conditions, or what severity of conditions should be used to define multi-morbidity.[9,28-30] Our results suggest that the underlying trait of multi-morbidity may be well captured by a multi-morbidity score with as few as 18 to 20 conditions, particularly at older ages. Furthermore, our results suggest that agreeing upon a universal set of conditions to define multi-morbidity may not be necessary if the conditions included in a score cover a broad range of body organ and sensory systems. Our findings are consistent with some models in frailty research which suggest that an accumulation of "deficits" is a proxy measure of aging.[31,32] The 4 multi-morbidity scores included in our study can be thought of as alternative ways of summarizing chronic conditions as accumulated deficits. Indeed, our previous results for mortality risk using percentile ranks of multi-morbidity (DHHS 20 conditions) are strikingly similar to results from a study using a frailty index

Table 3. Agreement in 6 pairwise comparisons between the 4 multi-morbidity score percentile ranks by age in 5-year increments for both sexes combined and separately for women and men.*

Sex/Age	Number of birthdays	Score comparisons, ICC (95% confidence interval)					
		DHHS vs. CCS	DHHS vs. CHAP	DHHS vs. ELIX	CCS vs. CHAP	CCS vs. ELIX	CHAP vs. ELIX
Both sexes							
20	20,468	0.62 (0.53 - 0.69)	0.58 (0.50 - 0.64)	0.80 (0.80 - 0.81)	0.96 (0.96 - 0.96)	0.70 (0.62 - 0.76)	0.66 (0.59 - 0.72)
25	22,376	0.64 (0.60 - 0.68)	0.61 (0.57 - 0.64)	0.78 (0.78 - 0.79)	0.96 (0.96 - 0.96)	0.74 (0.68 - 0.78)	0.70 (0.65 - 0.75)
30	22,769	0.64 (0.56 - 0.70)	0.61 (0.53 - 0.67)	0.78 (0.77 - 0.78)	0.96 (0.96 - 0.96)	0.72 (0.65 - 0.78)	0.69 (0.62 - 0.74)
35	19,534	0.67 (0.59 - 0.73)	0.64 (0.56 - 0.70)	0.78 (0.77 - 0.79)	0.96 (0.96 - 0.96)	0.72 (0.63 - 0.78)	0.69 (0.61 - 0.75)
40	17,733	0.71 (0.64 - 0.77)	0.68 (0.62 - 0.73)	0.77 (0.76 - 0.77)	0.96 (0.96 - 0.96)	0.74 (0.65 - 0.80)	0.71 (0.63 - 0.76)
45	19,366	0.75 (0.68 - 0.79)	0.71 (0.66 - 0.76)	0.77 (0.76 - 0.77)	0.96 (0.96 - 0.96)	0.74 (0.65 - 0.80)	0.70 (0.63 - 0.76)
50	21,151	0.78 (0.74 - 0.82)	0.75 (0.72 - 0.78)	0.77 (0.76 - 0.78)	0.96 (0.95 - 0.96)	0.75 (0.66 - 0.80)	0.71 (0.64 - 0.76)
55	19,272	0.82 (0.78 - 0.84)	0.78 (0.76 - 0.80)	0.78 (0.77 - 0.79)	0.95 (0.95 - 0.95)	0.76 (0.68 - 0.82)	0.72 (0.67 - 0.76)
60	15,218	0.84 (0.82 - 0.86)	0.79 (0.78 - 0.80)	0.78 (0.77 - 0.80)	0.94 (0.94 - 0.94)	0.77 (0.71 - 0.82)	0.72 (0.68 - 0.75)
65	11,870	0.85 (0.83 - 0.87)	0.80 (0.79 - 0.81)	0.80 (0.79 - 0.81)	0.93 (0.93 - 0.94)	0.79 (0.74 - 0.82)	0.73 (0.71 - 0.75)
70	9,420	0.83 (0.82 - 0.85)	0.76 (0.75 - 0.77)	0.79 (0.79 - 0.80)	0.90 (0.89 - 0.91)	0.79 (0.75 - 0.82)	0.69 (0.68 - 0.71)
75	7,356	0.83 (0.81 - 0.84)	0.73 (0.72 - 0.74)	0.79 (0.79 - 0.80)	0.88 (0.87 - 0.89)	0.80 (0.77 - 0.82)	0.68 (0.67 - 0.69)
80	5,876	0.83 (0.82 - 0.85)	0.72 (0.71 - 0.73)	0.79 (0.78 - 0.80)	0.86 (0.84 - 0.87)	0.80 (0.78 - 0.82)	0.67 (0.65 - 0.68)
85	4,281	0.84 (0.82 - 0.85)	0.72 (0.70 - 0.73)	0.79 (0.77 - 0.80)	0.84 (0.82 - 0.86)	0.82 (0.80 - 0.83)	0.68 (0.66 - 0.70)
All 18-85	1,075,433	0.74 (0.69 - 0.78)	0.70 (0.66 - 0.73)	0.79 (0.79 - 0.79)	0.95 (0.95 - 0.95)	0.75 (0.68 - 0.80)	0.71 (0.66 - 0.75)
Women							
20	10,711	0.63 (0.51 - 0.71)	0.58 (0.47 - 0.66)	0.81 (0.80 - 0.81)	0.96 (0.96 - 0.96)	0.71 (0.62 - 0.77)	0.67 (0.59 - 0.73)
25	12,572	0.63 (0.57 - 0.69)	0.60 (0.53 - 0.65)	0.79 (0.78 - 0.79)	0.96 (0.96 - 0.96)	0.73 (0.66 - 0.78)	0.70 (0.63 - 0.75)
30	12,097	0.60 (0.49 - 0.69)	0.57 (0.46 - 0.64)	0.76 (0.75 - 0.76)	0.96 (0.96 - 0.96)	0.72 (0.63 - 0.78)	0.69 (0.61 - 0.75)
35	9,988	0.64 (0.52 - 0.72)	0.60 (0.50 - 0.68)	0.77 (0.76 - 0.77)	0.96 (0.95 - 0.96)	0.73 (0.64 - 0.79)	0.70 (0.62 - 0.76)
40	9,076	0.67 (0.56 - 0.75)	0.64 (0.55 - 0.71)	0.77 (0.76 - 0.78)	0.95 (0.95 - 0.96)	0.75 (0.66 - 0.81)	0.72 (0.65 - 0.77)
45	10,015	0.70 (0.61 - 0.77)	0.66 (0.59 - 0.72)	0.77 (0.76 - 0.78)	0.95 (0.95 - 0.95)	0.74 (0.65 - 0.79)	0.70 (0.63 - 0.75)
50	11,034	0.74 (0.67 - 0.79)	0.71 (0.66 - 0.75)	0.76 (0.76 - 0.77)	0.95 (0.95 - 0.95)	0.75 (0.67 - 0.80)	0.71 (0.65 - 0.75)
55	10,065	0.79 (0.74 - 0.82)	0.75 (0.72 - 0.78)	0.77 (0.77 - 0.78)	0.94 (0.94 - 0.95)	0.76 (0.69 - 0.81)	0.72 (0.67 - 0.76)
60	7,943	0.82 (0.80 - 0.84)	0.78 (0.76 - 0.79)	0.78 (0.77 - 0.79)	0.94 (0.93 - 0.94)	0.77 (0.72 - 0.82)	0.72 (0.69 - 0.75)
65	6,286	0.84 (0.82 - 0.86)	0.80 (0.79 - 0.81)	0.80 (0.79 - 0.81)	0.93 (0.92 - 0.93)	0.78 (0.73 - 0.82)	0.73 (0.70 - 0.75)
70	4,981	0.81 (0.80 - 0.83)	0.75 (0.74 - 0.76)	0.78 (0.77 - 0.80)	0.90 (0.89 - 0.91)	0.77 (0.73 - 0.81)	0.68 (0.66 - 0.70)
75	3,979	0.83 (0.81 - 0.85)	0.75 (0.73 - 0.76)	0.79 (0.77 - 0.80)	0.89 (0.87 - 0.90)	0.79 (0.75 - 0.81)	0.69 (0.68 - 0.71)
80	3,298	0.83 (0.81 - 0.85)	0.74 (0.72 - 0.75)	0.79 (0.78 - 0.80)	0.86 (0.84 - 0.88)	0.79 (0.77 - 0.81)	0.68 (0.66 - 0.70)
85	2,611	0.83 (0.82 - 0.85)	0.72 (0.70 - 0.74)	0.78 (0.76 - 0.79)	0.85 (0.82 - 0.86)	0.81 (0.79 - 0.83)	0.69 (0.67 - 0.71)
All 18-85	568,558	0.71 (0.64 - 0.76)	0.67 (0.62 - 0.71)	0.79 (0.79 - 0.79)	0.95 (0.95 - 0.95)	0.75 (0.68 - 0.80)	0.71 (0.66 - 0.75)
Men							
20	9,757	0.61 (0.55 - 0.67)	0.57 (0.52 - 0.62)	0.79 (0.79 - 0.80)	0.95 (0.95 - 0.96)	0.69 (0.61 - 0.75)	0.65 (0.58 - 0.70)
25	9,804	0.65 (0.63 - 0.66)	0.61 (0.59 - 0.63)	0.77 (0.76 - 0.78)	0.95 (0.95 - 0.96)	0.74 (0.70 - 0.77)	0.70 (0.66 - 0.73)
30	10,672	0.68 (0.63 - 0.72)	0.65 (0.60 - 0.69)	0.80 (0.79 - 0.80)	0.96 (0.96 - 0.97)	0.71 (0.65 - 0.76)	0.68 (0.62 - 0.73)
35	9,546	0.70 (0.64 - 0.74)	0.67 (0.61 - 0.71)	0.79 (0.78 - 0.80)	0.97 (0.96 - 0.97)	0.70 (0.61 - 0.76)	0.66 (0.58 - 0.73)
40	8,657	0.75 (0.71 - 0.79)	0.72 (0.68 - 0.76)	0.76 (0.75 - 0.77)	0.97 (0.96 - 0.97)	0.72 (0.62 - 0.78)	0.69 (0.60 - 0.75)
45	9,351	0.79 (0.75 - 0.82)	0.76 (0.73 - 0.79)	0.77 (0.75 - 0.78)	0.97 (0.96 - 0.97)	0.73 (0.63 - 0.80)	0.70 (0.62 - 0.76)
50	10,117	0.82 (0.80 - 0.85)	0.79 (0.77 - 0.81)	0.77 (0.75 - 0.79)	0.96 (0.96 - 0.97)	0.75 (0.65 - 0.81)	0.71 (0.63 - 0.76)
55	9,207	0.85 (0.83 - 0.87)	0.82 (0.81 - 0.83)	0.78 (0.76 - 0.80)	0.96 (0.95 - 0.96)	0.77 (0.67 - 0.83)	0.73 (0.66 - 0.77)
60	7,275	0.86 (0.84 - 0.87)	0.81 (0.80 - 0.82)	0.79 (0.76 - 0.81)	0.94 (0.94 - 0.95)	0.77 (0.70 - 0.82)	0.71 (0.67 - 0.75)
65	5,584	0.86 (0.85 - 0.88)	0.80 (0.79 - 0.81)	0.81 (0.80 - 0.82)	0.93 (0.93 - 0.94)	0.80 (0.75 - 0.83)	0.73 (0.71 - 0.75)
70	4,439	0.85 (0.84 - 0.87)	0.77 (0.75 - 0.78)	0.81 (0.79 - 0.82)	0.90 (0.89 - 0.91)	0.80 (0.76 - 0.83)	0.70 (0.69 - 0.72)
75	3,377	0.83 (0.81 - 0.84)	0.70 (0.69 - 0.72)	0.81 (0.79 - 0.82)	0.88 (0.86 - 0.89)	0.81 (0.78 - 0.83)	0.66 (0.65 - 0.68)
80	2,578	0.83 (0.81 - 0.85)	0.70 (0.68 - 0.72)	0.80 (0.78 - 0.81)	0.85 (0.83 - 0.86)	0.80 (0.78 - 0.82)	0.65 (0.63 - 0.67)
85	1,670	0.84 (0.82 - 0.86)	0.71 (0.69 - 0.74)	0.80 (0.78 - 0.82)	0.84 (0.82 - 0.86)	0.83 (0.80 - 0.84)	0.67 (0.64 - 0.69)
All 18-85	506,875	0.77 (0.74 - 0.79)	0.73 (0.71 - 0.75)	0.80 (0.79 - 0.81)	0.95 (0.95 - 0.96)	0.75 (0.68 - 0.80)	0.71 (0.65 - 0.75)

ICC = intraclass correlation coefficient; DHHS = Department of Health and Human Services, 20 condition score; CCS = Clinical Classifications Software, 190 chronic condition score; CHAP = Body system chapters (collapsed CCS system), 18 chapters score; ELIX = Elixhauser comorbidity index, 31 condition score.
*The persons who experienced multiple birthdays during the study period were included multiple times (at different ages).

that included 46 deficits.[10,33] Therefore, the current study extends to multi-morbidity research a methodology that has been used previously in some frailty research.[34]

Strengths and weaknesses

The strengths of this study include access to medical record data (specifically, billing codes) on all conditions for an entire geographically-defined population across age, sex, race, and ethnic groups, regardless of insurance status, socioeconomic status, and care delivery setting. This study design should have minimized selection bias. Medical record data are often difficult to obtain in the United States because there is no centralized health care surveillance system and no comprehensive, clinical records-based data sets for all ages and for all regions.[35] Therefore, the REP offered a unique setting to study multi-morbidity across all ages within a local context.[5,11] In addition, because data were generated historically as part of routine medical care, persons were not involved in remembering or reporting medical events or diagnoses. This study design should have minimized recall bias.

A limitation of this study, which is shared with many other similar studies, is the limited validity of ICD-9 codes from administrative databases to define each specific chronic condition. Previous REP studies have shown that codes may be assigned in error, and manual review of the medical records is often needed to ascertain whether a person truly has the disease or condition of interest.[5,36-40] We recently published a comparison of ICD-9 codes vs. nurse abstraction of the full medical charts to identify 17 of the 20 chronic conditions included in the DHHS score.[41] The remaining 3 conditions could not be studied because they were too uncommon in the study sample (hepatitis, HIV infection, and autism spectrum disorder). Sensitivity and positive predictive value varied across the 17 chronic conditions, and were also influenced by the use of at least 1 diagnostic code or of at least 2 diagnostic codes separated by more than 30 days.[41] In this study, we limited false positive diagnoses by requiring at least 2 diagnostic codes separated by more than 30 days to define each individual condition for all 4 of the multi-morbidity scores. However, we may have undercounted certain conditions (reduced sensitivity).

Second, we only considered 4 multi-morbidity measures. There is not agreement on the optimal number and type of conditions needed to define multi-morbidity, or on the advantages and disadvantages of using weights for the conditions included. The optimal way to measure multi-morbidity likely depends on the purpose of the study.[28-30] For example, we have consistently used the DHHS score in our previous studies of multi-morbidity.[2,5,6,9,10,41] In this study, we investigated 4 multi-morbidity scores, all of which are defined using readily accessible ICD-9 code listings.[15-18,20] We calculated simple unweighted multi-morbidity scores of the number of conditions at fixed ages. Because of feasibility and space limitations, we did not include all of the commonly used multi-morbidity scores in our comparisons.[29] However, we included scores with a wide range in number and type of included conditions (from a condensed grouping of 18 body-system chapters to 190 individual chronic conditions). In addition, we did not compare the 4 multi-morbidity scores in their ability to predict specific outcomes (e.g., hospitalization or 1-year mortality). Our analyses focused on whether the 4 multi-morbidity scores agreed in their measurement of the underlying "construct" of multi-morbidity in a well-defined population.

A third limitation of our study is the cross-sectional nature of the analyses used to develop the percentile ranks.[42] The persons residing in the county over a 10 calendar-year period were sampled based on having reached a certain birthday age. Therefore, the percentile distributions were based on the assumptions that all persons in a given age and sex stratum had the same multi-morbidity score profile regardless of the calendar year of the measure or of the calendar year of their birth. Our previous study of the DHHS score percentiles showed no systematic shifts in the distribution of multi-morbidity scores over the 10 calendar-year study period and no birth cohort effects.[10]

Fourth, we did not assign percentile ranks separately in strata by race and ethnicity. Indeed, our previous study found that the percentile distribution of the DHHS scores did not vary by race or ethnicity in our population.[10] In the race and ethnicity stratified analyses in this study, the differences in agreement of the 4 multi-morbidity scores were small.

Finally, our study focused on a single geographically-defined US population, and the percentile distribution of multi-morbidity may differ in other populations. Nevertheless, the demographic and socioeconomic characteristics of our population are similar to those of the upper Midwest and of a large segment of the entire US population.[27] In addition, we emphasize that for our analyses of the agreement of the 4 score percentile ranks, a person functioned as his/her own comparison (i.e., 4 different scores were calculated using the same diagnostic information in a 5-year window of time before an index birthday). However, the percentile ranks may exhibit lower agreement in other populations if the frequencies of underlying conditions that make up the scores vary substantially. Replication of this study in other populations in the US and worldwide will allow for a more thorough understanding of the agreement between multi-morbidity measures.
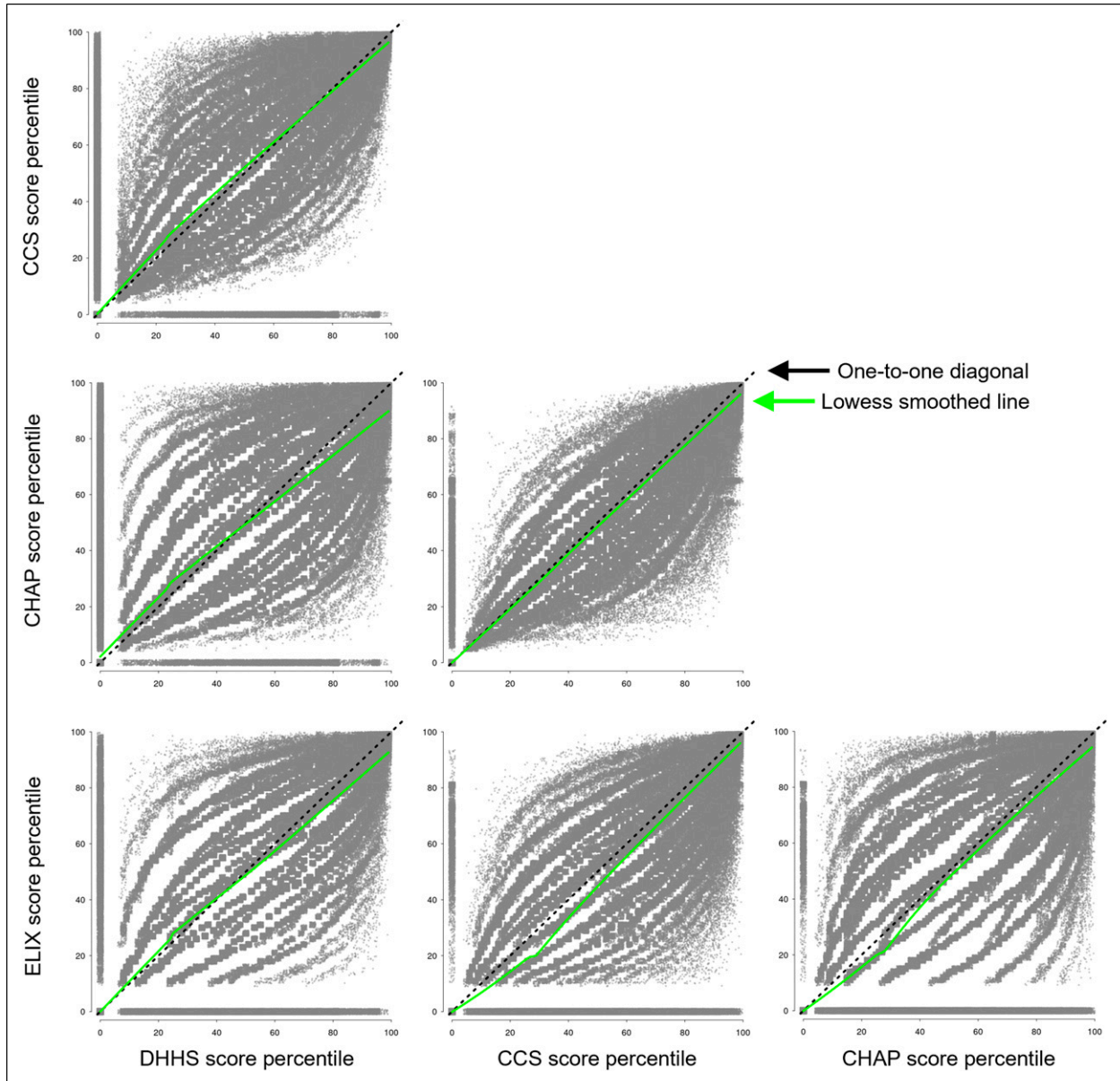
Figure 1. Quantile-quantile scatter plots of percentile ranks for 6 pairwise comparisons between the 4 multi-morbidity scores. For visual reference, the one-to-one diagonal line (black dashed line and black arrow) and the lowess smoothed line (green solid line and green arrow) are shown. Strong agreement between the scores was observed for all 6 comparisons. Abbreviations: CCS = Clinical Classifications Software, 190 chronic condition score; CHAP = Body system chapters (collapsed CCS system), 18 chapters score; DHHS = Department of Health and Human Services, 20 condition score; ELIX = Elixhauser comorbidity index, 31 condition score.

Comparison with other studies

There is a growing body of literature describing the patterns of multi-morbidity in populations in the United States and worldwide;[28,29] however, we are not aware of studies that have compared percentile ranks across several multi-morbidity scores to assess agreement. Our findings suggest that the use of percentile ranks is consistent across 4 multi-morbidity scores applied to the same population. Further studies are needed to confirm that this convergence of measures is also maintained across populations with different baseline prevalence of conditions. If this convergence is maintained, results from studies using different numbers or types of chronic conditions to define multi-morbidity may be compared using percentile ranks. Our results also suggest that a limited set of conditions covering
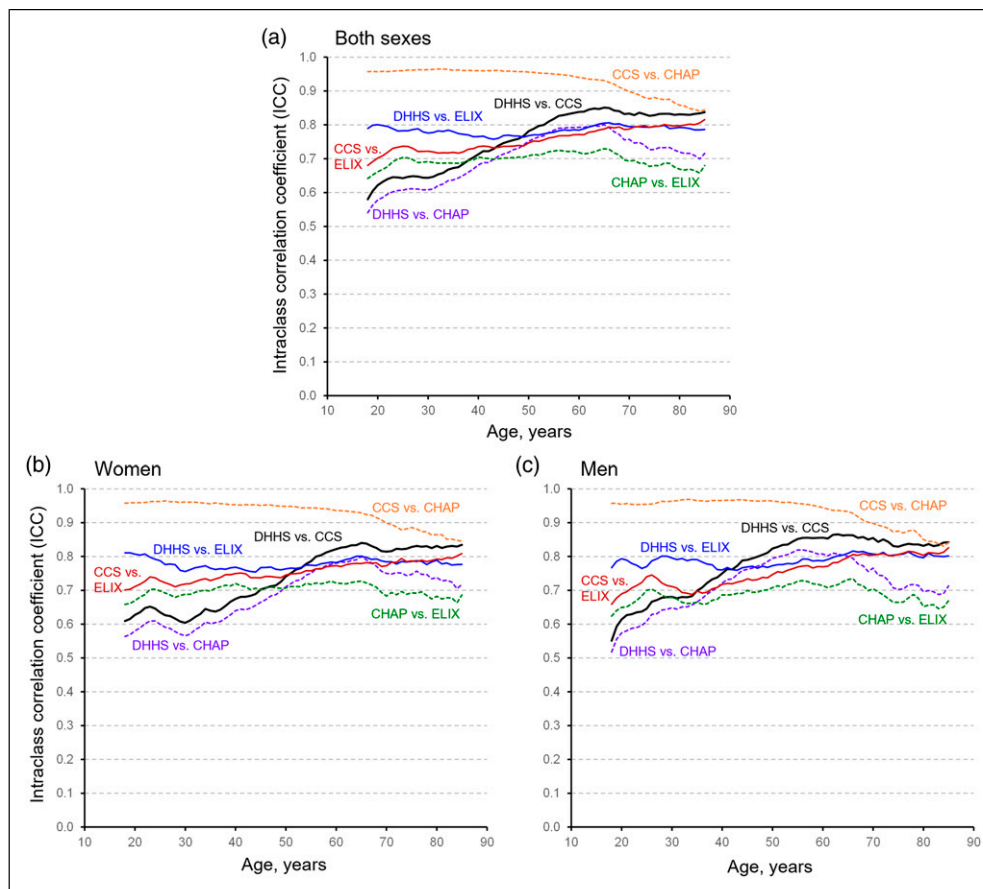
Figure 2. Intraclass correlation coefficients for 6 pairwise comparisons between the 4 multi-morbidity percentile ranks plotted across age (18 to 85 years) for both sexes combined (panel A) and separately for women (panel B) and men (panel C). Abbreviations: CCS = Clinical Classifications Software, 190 chronic condition score; CHAP = Body system chapters (collapsed CCS system), 18 chapters score; DHHS = Department of Health and Human Services, 20 condition score; ELIX = Elixhauser comorbidity index, 31 condition score.

the body's key sensory and organ systems may be sufficient to capture multi-morbidity status, particularly at ages 50 years and older. This finding has both clinical and research implications. Further research is needed to determine whether percentile ranks may be a more comparable way of predicting future adverse outcomes (e.g., hospitalization or short-term mortality) in the absence of an agreed upon universal measure of multi-morbidity.

Conclusions

Percentile ranks may provide a simple and intuitive measure of the health of persons as compared to peers of the same age and same sex living in the same population. The use of percentile ranks in research projects may also increase comparability across populations for studies using different numbers or types of chronic conditions to define multi-morbidity.

List of abbreviations

CCS: Clinical Classifications Software (CSS)
CHAP: CCS Chapters
DHHS: US Department of Health and Human Services
ELIX: Elixhauser comorbidity index
gcrq: growth charts regression quantiles
HIV: human immunodeficiency virus
ICC: Intra-class correlation coefficients
ICD-9: International Classification of Diseases, Ninth Revision
irr: inter-rater reliability
REP: Rochester Epidemiology Project

Author contributions

BRG, JLS, and WAR were involved in the conception and design of the study. BRG and WAR conducted the data analyses. BRG drafted the manuscript. All authors (BRG, AMC, CMB, WVB, JLS, and WAR) contributed to the interpretation of the data and provided critical revisions of the manuscript. All authors (BRG, AMC, CMB, WVB, JLS, and WAR) also approved the final version to be published.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Ethics approval

All study procedures and ethical aspects were approved by the institutional review boards of both Mayo Clinic and Olmsted Medical Center. Because the data collection was historical, persons were not required to provide a study-specific informed consent but did provide a general consent to use their medical records for research (Minnesota legal requirements).

Data availability

To obtain the data that support the findings of this study, the user must submit a brief outline of the intended use of the data (no longer than a page). In addition, the user will be required to sign a data sharing agreement and pay for the preparation of the shared file of de-identified data.

Disclaimer

The content of this article is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health (NIH) or the Mayo Clinic.

ORCID iDs

Alanna M. Chamberlain https://orcid.org/0000-0003-1888-9584
William V. Bobo https://orcid.org/0000-0002-5674-3079
Jennifer L. St Sauver https://orcid.org/0000-0002-9789-8544
Walter A. Rocca https://orcid.org/0000-0002-1832-7664

Supplemental Material

Supplemental material for this article is available online.

References

1. Fabbri E, Zoli M, Gonzalez-Freire M, et al. Aging and multimorbidity: new tasks, priorities, and frontiers for integrated gerontological and clinical research. *J Am Med Dir Assoc* 2015; 16: 640–647.

2. St Sauver JL, Boyd CM, Grossardt BR, et al. Risk of developing multimorbidity across all ages in an historical cohort study: differences by sex and ethnicity. *BMJ Open* 2015; 5: e006413.

3. Fabbri E, An Y, Zoli M, et al. Association between accelerated multimorbidity and age-related cognitive decline in older Baltimore longitudinal study of aging participants without dementia. *J Am Geriatr Soc* 2016; 64: 965–972.

4. Vetrano DL, Calderon-Larranaga A, Marengoni A, et al. An International Perspective on Chronic Multimorbidity: Approaching the Elephant in the Room. *J Gerontol A Biol Sci Med Sci* 2018; 73: 1350–1356.

5. Rocca WA, Boyd CM, Grossardt BR, et al. Prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and race/ethnicity. *Mayo Clin Proc* 2014; 89: 1336–1349.

6. Bobo WV, Yawn BP, St Sauver JL, et al. Prevalence of Combined Somatic and Mental Health Multimorbidity: Patterns by Age, Sex, and Race/Ethnicity. *J Gerontol A Biol Sci Med Sci* 2016; 71: 1483–1491.

7. Violan C, Foguet-Boreu Q, Roso-Llorach A, et al. Burden of multimorbidity, socioeconomic status and use of health services across stages of life in urban areas: a cross-sectional study. *BMC Public Health* 2014; 14: 530.

8. Barnett K, Mercer SW, Norbury M, et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012; 380: 37–43.

9. Rocca WA, Gazzuola-Rocca L, Smith CY, et al. Accelerated accumulation of multimorbidity after bilateral oophorectomy: a population-based cohort study. *Mayo Clin Proc* 2016; 91: 1577–1589.

10. Rocca WA, Grossardt BR, Boyd CM, et al. Multimorbidity, ageing and mortality: normative data and cohort study in an American population. *BMJ Open* 2021; 11: e042633.

11. Posner SF and Goodman RA. Multimorbidity at the local level: Implications and research directions. *Mayo Clin Proc* 2014; 89: 1321–1323.

12. Shrout PE and Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86: 420–428.

13. McGraw KO and Wong SP. Forming inferences about some intraclass correlation coefficients. 1996; 1: 30–46.

14. St Sauver JL, Grossardt BR, Yawn BP, et al. Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. *Int J Epidemiol* 2012; 41: 1614–1624.

15. Goodman RA, Posner SF, Huang ES, et al. Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. *Prev Chronic Dis* 2013; 10: E66.

16. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP) Clinical Classifications Software (CCS) for ICD-9-CM. https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. (2012, accessed November 14, 2019).

17. Agency for Healthcare Research and Quality. Medical Expenditure Panel Survey HC-120, Appendix 3: Clinical Classification Code to ICD-9-CM Code Crosswalk. https://meps.ahrq.gov/data_stats/download_data/pufs/h120/h120app3.html (accessed November 14, 2019).

18. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP) Chronic Condition Indicator (CCI). https://www.hcup-us.ahrq.gov/toolssoftware/chronic/chronic.jsp (2016, accessed May 2021).

19. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. *Med Care* 1998; 36: 8–27.

20. 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: 1130–1139.

21. Garland A, Fransoo R, Olafson K, et al. *The Epidemiology and Outcomes of Critical Illness in Manitoba*. April 2012 2012. Winnipeg, MB: Manitoba Centre for Health Policy.

22. Sharabiani MT, Aylin P and Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care* 2012; 50: 1109–1118.

23. National Center for Health Statistics. *The International Classification of Diseases, 9th revision, Clinical Modification: ICD-9-CM*. Ann Arbor, Mich: Commission on Professional and Hospital Activities, 1980.

24. Machado JAF and Silva JMCS. Quantiles for Counts. *Journal of the American Statistical Association* 2005; 100: 1226–1237.

25. Gamer M, Lemon J, Fellows I, et al. *Package 'irr': Various Coefficients of Interrater Reliability and Agreement*. https://cran.r-project.org/web/packages/irr/ (2019, accessed May 2021).

26. Cleveland WS. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 1979; 74: 829–836.

27. St Sauver JL, Grossardt BR, Leibson CL, et al. Generalizability of epidemiological findings and public health decisions: an illustration from the Rochester Epidemiology Project. *Mayo Clin Proc* 2012; 87: 151–160.

28. Salive ME, Suls J, Farhat T, et al. National Institutes of Health Advancing Multimorbidity Research. *Med Care* 2021; 59: 622–624.

29. Suls J, Bayliss EA, Berry J, et al. Measuring Multimorbidity: Selecting the Right Instrument for the Purpose and the Data Source. *Med Care* 2021; 59: 743–756.

30. The Academy of Medical Sciences. Multimorbidity: a priority for global health research. https://acmedsci.ac.uk/policy/policy-projects/multimorbidity (2018, accessed May 2021).

31. Mitnitski AB, Mogilner AJ and Rockwood K. Accumulation of deficits as a proxy measure of aging. *ScientificWorldJournal* 2001; 1: 323–336.

32. Rockwood K. What would make a definition of frailty successful? *Age Ageing* 2005; 34: 432–434.

33. Rockwood K, Blodgett JM, Theou O, et al. A Frailty Index Based On Deficit Accumulation Quantifies Mortality Risk in Humans and in Mice. *Sci Rep* 2017; 7: 43068.

34. Rutenberg AD, Mitnitski AB, Farrell SG, et al. Unifying aging and frailty through complex dynamical networks. *Exp Gerontol* 2018; 107: 126–129.

35. Wallace RB and Salive ME. The dimensions of multiple chronic conditions: where do we go from here? A commentary on the Special Issue of Preventing Chronic Disease. *Prev Chronic Dis* 2013; 10: E59.

36. Leibson CL, Brown AW, Ransom JE, et al. Incidence of traumatic brain injury across the full disease spectrum: a population-based medical record review study. *Epidemiology* 2011; 22: 836–844.

37. Leibson CL, Naessens JM, Brown RD Jr., et al. Accuracy of hospital discharge abstracts for identifying stroke. *Stroke* 1994; 25: 2348–2355.

38. Leibson CL, Needleman J, Buerhaus P, et al. Identifying in-hospital venous thromboembolism (VTE): a comparison of claims-based approaches with the Rochester Epidemiology Project VTE cohort. *Med Care* 2008; 46: 127–132.

39. Roger VL, Killian J, Henkel M, et al. Coronary disease surveillance in Olmsted County objectives and methodology. *J Clin Epidemiol* 2002; 55: 593–601.

40. Yawn BP, Wollan P and St Sauver J. Comparing shingles incidence and complication rates from medical record review and administrative database estimates: how close are they? *Am J Epidemiol* 2011; 174: 1054–1061.

41. St Sauver JL, Chamberlain AM, Bobo WV, et al. Implementing the US Department of Health and Human Services definition of multimorbidity: a comparison between billing codes and medical record review in a population-based sample of persons 40-84 years old. *BMJ Open* 2021; 11: e042870.

42. O'Connor PJ. Normative data: their definition, interpretation, and importance for primary care physicians. *Fam Med* 1990; 22: 307–311.