ORIGINAL RESEARCH

Charlson comorbidity index derived from chart review or administrative data: agreement and prediction of mortality in intensive care patients

Knut Stavem[1–3]
Henrik Hoel[4]
Stein Arve Skjaker[5]
Rolf Haagensen[6]

[1]Division of Medicine and Laboratory Sciences (AHUSKIL), Campus Ahus, Institute of Clinical Medicine, University of Oslo, Oslo, [2]Department of Pulmonary Medicine, Medical Division, [3]Health Services Research Unit, Akershus University Hospital, Lørenskog, [4]Department of Surgery, Sykehuset Innlandet Kongsvinger, Kongsvinger, [5]Section of Orthopaedic Emergency, Department of Orthopaedic Surgery, Oslo University Hospital, Oslo, [6]Department of Anaesthesiology, Surgical Division, Akershus University Hospital, Lørenskog, Norway

Correspondence: Knut Stavem
Health Services Research Unit, Akershus University Hospital, Sykehusveien 25, N-1478 Lørenskog, Norway
Tel +47 9754 2157
Email knut.stavem@medisin.uio.no

Purpose: This study compared the Charlson comorbidity index (CCI) information derived from chart review and administrative systems to assess the completeness and agreement between scores, evaluate the capacity to predict 30-day and 1-year mortality in intensive care unit (ICU) patients, and compare the predictive capacity with that of the Simplified Acute Physiology Score (SAPS) II model.

Patients and methods: Using data from 959 patients admitted to a general ICU in a Norwegian university hospital from 2007 to 2009, we compared the CCI score derived from chart review and administrative systems. Agreement was assessed using % agreement, kappa, and weighted kappa. The capacity to predict 30-day and 1-year mortality was assessed using logistic regression, model discrimination with the c-statistic, and calibration with a goodness-of-fit statistic.

Results: The CCI was complete (n=959) when calculated from chart review, but less complete from administrative data (n=839). Agreement was good, with a weighted kappa of 0.667 (95% confidence interval: 0.596–0.714). The c-statistics for categorized CCI scores from charts and administrative data were similar in the model that included age, sex, and type of admission: 0.755 and 0.743 for 30-day mortality, respectively, and 0.783 and 0.775, respectively, for 1-year mortality. Goodness-of-fit statistics supported the model fit.

Conclusion: The CCI scores from chart review and administrative data showed good agreement and predicted 30-day and 1-year mortality in ICU patients. CCI combined with age, sex, and type of admission predicted mortality almost as well as the physiology-based SAPS II.

Keywords: epidemiology, Charlson comorbidity index, intensive care unit, comorbidity, mortality, SAPS II, case-mix adjustment, agreement, calibration, discrimination, prediction

Introduction

Administrative data are often used for risk adjustment in observational studies of variations in utilization, expenditure, and mortality rates, for example in selected diseases, across hospitals, regions, or countries. In such observational studies, it is essential to properly adjust for differences in case mix, eg, through the use of measures of disease severity or comorbidity. In intensive care patients, physiology-based severity-of-illness scores can be used to adjust for patient demographics, comorbidities, and illness severity.[1,2] These measures, however, require additional prospective data collection, are complex, are not readily available in administrative databases, and are intended for prediction of in-hospital mortality after admission to an intensive care unit (ICU). Comorbidity is also an important determinant of the outcome of an ICU stay and may improve prognostic predictions for critically ill patients.[3] The measurement of comorbidity has received increasing attention in the literature, though no consensus regarding the definition of comorbidity currently exists.[4]

The Charlson comorbidity index (CCI) is among the best-known and widely used indexes of comorbidity.[4–8] Data for the CCI can be obtained from chart review[5] or derived from administrative data with ICD-9CM or ICD-10 codes.[9–11] Chart review gives a higher prevalence of comorbidities than administrative data,[12,13] though agreement between data sources varies according to patient population. However, this finding is not universal, and another study reported higher CCI scores and better discriminative capacity of CCI when using administrative data than when using single-day chart review in hospitalized patients.[14] Furthermore, chart review can be very time consuming, whereas administrative databases can yield data from large patient groups.[4]

Several studies have documented the validity of using the CCI scores obtained from chart data[3] and administrative data[15–18] to predict mortality in ICU patients. Moreover, critical care populations vary considerably between studies; hence, there is a need to validate the use of the CCI in a variety of populations. Few studies have compared CCI scores from different sources and validated their respective predictive capacities.[19] Furthermore, CCI scores derived from ICD-9CM codes and ICD-10 codes are not necessarily equivalent, and coding practices may differ both locally and between countries.

The objectives of this study of 959 intensive care patients from a general ICU were to compare the following: 1) the completeness and agreement between CCI scores derived from chart review and administrative data for prior hospital contacts (using ICD-10 codes); 2) the capacity of the derived CCI scores to predict 30-day and 1-year mortality; and 3) the predictive capacity with that of the Simplified Acute Physiology Score II (SAPS II)[1] model.

Patients and methods
Patient population and study design
The study was an observational study of ICU patients from Akershus University Hospital, a 615-bed general hospital that covers all medical and surgical specialties except neurosurgery, multitrauma care, cardiac surgery, eye diseases, and rheumatology. It is affiliated with the University of Oslo. All patients with acute conditions in the catchment area are treated in the hospital, although for elective conditions, every patient – by law – has the opportunity to choose the hospital for treatment. However, few patients in practice take advantage of this legal right. For specialties that are not covered, the hospital refers patients to other tertiary care hospitals in the area.

At the time of the study, the hospital had a catchment population of ~320,000 people in Oslo and Akershus. The hospital had a 9-bed general ICU and had treated patients from the medical, surgical, and neurological departments. From January 1, 2007 to December 31, 2009, this ICU had 1,287 patient admissions. After excluding 328 patients with age <18 years, ICU stay <24 hours, or readmission to the ICU, we included 959 patients with first admission to the ICU (termed "index admission") during this 3-year period.[20]

The study was presented to the Regional Committee for Medical Health Research Ethics and was determined to be a quality assurance study (ref: 2011/950). It was approved by the local privacy ombudsman for research at Akershus University Hospital. Patient consent was not required to access and review the data.

Data sources and data collection
Chart review
The electronic medical charts of all patients were reviewed by one of the authors (HH or SAS), who recorded the patient's comorbidity score according to CCI.[5] The reviewer used all available information in the medical record until admission to the ICU, including running text in admission reports, case notes, discharge summaries, and laboratory data back to 2004, and – when available – scanned versions of earlier paper-based medical records or scanned discharge summaries or case notes from other hospitals. The presence of the medical conditions of the CCI was marked on a registration form.

Patient administrative data
Date and time of the index hospital admission, ICU admission, and ICU discharge were extracted from the hospital's patient administration system to estimate the length of the hospital stay prior to ICU admission.

For the purpose of scoring CCI from administrative data, we extracted each patient's ICD-10 codes from all previous inpatient admissions and hospital outpatient visits between January 1, 1999, when ICD-10 coding was introduced, and the index admission. In Norway, for reimbursement purposes, 1 primary diagnosis, a maximum of 7 secondary diagnoses, and 10 procedure codes are allowed for each admission or outpatient visit. However, most outpatient visits are coded with only 1–2 diagnoses. We defined comorbidity as conditions present upon admission. Hence, we only selected inpatient admissions and ambulatory visits occurring prior to the index admission, as in a previous study.[12]

CCI scoring

The CCI was originally developed to predict 1-year mortality in a mixed population of internal medicine patients using comorbidity derived from chart review.[5] The CCI consists of 19 selected conditions that are weighted and summed to an index on a 0–33 scale. Its reliability and validity have been assessed in various ways and in different populations, and in general, its test–retest and interrater reliability is moderate to good.[4,21] More recently, this index has been adapted for use with diagnoses from administrative databases, and revised weights have been suggested or adapted for specific populations.[9,22,23] The CCI is typically used to describe populations, predict outcomes, adjust for case mix to reduce confounding in observational studies in conjunction with a variety of different outcomes, such as mortality, disability, readmissions, and length of stay.

For the CCI scores obtained from the electronic charts, we used the weights from the original CCI, producing a score using a 0–33 scale.[5] For the CCI scores obtained from the administrative data, CCI scores were derived using a previously validated algorithm that is based on ICD-10 data; it scores the CCI on a 0–29 scale.[10,23] When no ICD-10 code was available from a previous hospital contact (ie, a missing value), the CCI was scored as 0.

Physiology-based severity score

During the first 24 hours of their ICU stay, patients were prospectively scored according to the SAPS II.[2] The SAPS II is a severity score for ICU patients and is based on the worst values measured on 17 variables during the initial 24 hours of their ICU stay. It is scored on a 0- to 163-point scale, with higher score indicating more severe illness. In addition, a mathematical formula is provided for predicting hospital mortality.

The SAPS II scores were stored in a local database, together with information on age, sex, type of admission, outcome of stay, time on mechanical ventilation, and if treatment was withheld or withdrawn. We defined "type of admission" as defined according to SAPS II[2]: medical (ie, patients with no surgery in the last 7 days before ICU admission), scheduled surgical (ie, patients whose surgery was scheduled at least 24 hours in advance), or unscheduled surgical (ie, patients added to operating room schedule within 24 hours of their operation).

Other data and mortality

Patients were also prospectively grouped into 8 principal disease categories that were based on the major disease categories in the Acute Physiology and Chronic Health Evaluation III (APACHE III) prognostic system: respiratory, cardiovascular/vascular, gastrointestinal, neurological, sepsis-related, trauma-related, metabolic, and other diseases.[24] We used this classification for descriptive purposes only.

We identified mortality cases during the study period through linkage to the registers of Statistics Norway until June 21, 2013, using the patients' unique national 11-digit person identification number.

Statistical analysis

We present descriptive statistics using the mean (standard deviation [SD]), the median (25th–75th percentiles), or count (%). Groups were compared using the independent samples t-test for normally distributed data or the Mann–Whitney U-test for nonnormally distributed data. Categorical data were compared using the chi-square test or Fisher's exact test, as appropriate.

Further analyses were conducted in the total study sample. CCI scores were categorized as 0, 1, and ≥2. The CCI score distributions obtained with the 2 calculation methods were compared using McNemar's test. Agreement between CCI from chart review and patient administration data was assessed using 1) Spearman's rank correlation between the scores, 2) the proportion of agreement between the categorized scores, and 3) weighted kappa with 95% confidence intervals (CIs) between categorized CCI scores from the 2 sources, using quadratic disagreement weights.

The individual items of the Charlson score were compared between the 2 data sources using 1) the proportion of agreement, 2) simple kappa with 95% CIs, and 3) positive predictive value (PPV) of the categorized CCI scores from administrative data, using the scores from the chart review as the reference. The kappa values were interpreted as follows: poor agreement (<0.20), fair agreement (0.21–0.40), moderate agreement (0.41–0.60), good agreement (0.61–0.80), and very good agreement (0.81–1.00).[25]

The predictive validity of the derived CCI scores was assessed using multivariate logistic regression analysis, with 30-day and 1-year mortality as the outcome measure. We assessed the discrimination of the models using the c-statistic, which represents the area under the receiver-operator characteristic curve. Model calibration was assessed using the Hosmer–Lemeshow goodness-of-fit statistic,[26] whereby we considered a value of $p<0.05$ to indicate that the model had a poor fit.

We assessed the impact of several combinations of covariates that were expected to be associated with mortality: age, sex, and type of admission (ie, medical, unscheduled

surgical, and scheduled surgical). We were also interested in determining whether the CCI score provided additive prognostic value to the SAPS II score in estimating mortality. In the regression models, we included only those variables that were available after the first 24 hours of ICU stay, without later updating of variables.

To check for multicollinearity issues, we assessed Spearman's rank correlation between the independent variables in the multivariable model and found that all pairwise correlations had values <0.45.

We used Stata version 14.1 for all analyses (Stata Inc, College Station, TX, USA). We chose a 5% significance level, using 2-sided tests.

Results

Of the 959 patients admitted to the ICU in the study period, 839 had information on previous admissions or visits that was available in the hospital's patient administration system. The 120 patients without information on previous admissions

or visits were younger, had lower CCI scores based on chart review, and experienced a shorter time between the hospital admission and transfer to the ICU than the 839 with available information; all other patient characteristics did not differ between the 2 groups (Table 1).

A scatter plot of the CCI scores showed that many of the patients were located away from the diagonal, suggesting a difference in the CCI scores derived from the 2 sources (Figure 1). The correlation between CCI scores derived from chart review and administrative data was 0.662 for scales categorized as 0, 1, and ≥2 and 0.692 for CCI scores on the continuous scales that were scored from 0 to 29 (n=33) (Table 2).

The agreement between CCI scores according to source, as assessed with weighted kappa with 95% CI (n=959), was 0.667 (0.596–0.714) for CCI administrative data vs CCI chart review (continuous); and 0.652 (0.613–0.697) for CCI administrative data vs CCI chart review (3 categories: 0, 1, and ≥2). The proportion of agreement between the categorized CCI

Table 1 Characteristics of intensive care patients included in the study (n=959) according to complete or incomplete scores

Characteristics	CCI from chart; missing administrative data score (n=120)	CCI score from chart and administrative data (n=839)	*p*-value
Age in years, mean (SD)	54.2 (18.4)	59.2 (18.1)	0.005
Sex, n (%)			0.703
Women	49 (41)	358 (43)	
Men	71 (59)	481 (57)	
Diagnostic categories, n (%)			0.228
Respiratory	29 (24)	225 (27)	
Cardiovascular	19 (16)	136 (16)	
Gastrointestinal	16 (13)	129 (15)	
Neurological	17 (14)	81 (10)	
Sepsis	6 (5)	66 (8)	
Trauma	12 (10)	40 (5)	
Metabolic	12 (10)	88 (10)	
Others	9 (8)	74 (9)	
Type of admission, n (%)			0.191
Medical	93 (78)	610 (73)	
Unscheduled surgical	25 (21)	184 (22)	
Scheduled surgical	2 (2)	45 (5)	
Ventilator treatment, n (%)			0.503
Yes	98 (82)	663 (79)	
No	22 (18)	176 (21)	
In-hospital discharge, n (%)			0.458
Dead	30 (25)	237 (28)	
Alive	90 (75)	602 (72)	
SAPS II score, mean (SD)	42.3 (17.8)	42.4 (17.1)	0.95
Charlson score (from charts), n (%)			<0.001
0	68 (57)	296 (35)	
1	31 (26)	170 (20)	
≥2	21 (18)	373 (45)	
Time from admission to ICU, days			<0.001
Admission, median (25th–75th percentiles)	0.09 (0.04-0.57)*	0.39 (0.07–3.06)**	

Notes: *n=119; **n=817.
Abbreviations: CCI, Charlson comorbidity index; ICU, intensive care unit; SAPS, Simplified Acute Physiology Score; SD, standard deviation.

scores from the 2 sources was 69% (657/959) (Table 3). The distributions of categorized CCI differed ($p<0.001$), with a larger proportion of patients with an index of 1 or ≥2 from chart review than from administrative data.

The agreement between CCI scores for individual comorbidities derived from the 2 data sources ranged from 88.8% to 99.6%, and kappa ranged from 0.26 to 0.66. PPVs, ie, the proportion of cases with a condition present in the chart review, given that the presence of the same condition based on the codes in the administrative data, ranged from 33.3% to 81.3% (Table 4).

In total, 249 (26%) patients died within 30 days and 341 (36%) patients died within 1 year. In the assessment of discrimination in logistic regression models for 30-day mortality, the c-statistic for the CCI score from chart review was 0.656 (95% CI: 0.621–0.692), which was somewhat higher than that of the administrative system-based score of 0.607 (95% CI: 0.570–0.644). The addition of age, sex, type of

admission, and ventilator status increased the c-statistic for models with the CCI from both sources, and the c-statistic values appeared to converge when adding additional variables into the models (Table 5). SAPS II scores had a higher c-statistic than the 2 CCI scores from different sources, and the numerical value was also higher than for the models with CCI score, age, sex, and type of admission. Furthermore, adding the CCI score to a model with the SAPS II score only marginally improved the discrimination of the model.

In models of 1-year mortality, the pattern of capacity to discriminate between patients was similar to that of the models of 30-day mortality. The CCI-based models performed somewhat better and the physiology-based models somewhat worse for 1-year mortality than for 30-day mortality (Table 5).

Model calibration showed satisfactory calibration ($p>0.05$) for most models, with the exception of SAPS II+CCI chart review for 30-day mortality (Table 6).

Discussion

In this study, the CCI score from chart review resulted in a higher CCI score than that derived from administrative systems, although the agreement between CCI scores derived from the 2 data sources was good and varied for the individual comorbidities. When 30-day mortality and 1-year mortality were set as outcome variables, the discrimination of the CCI combined with age, sex, and type of admission was likewise good, and model calibration was acceptable. In this ICU population, however, the physiology-based SAPS II score had better predictive properties, and there was little additional
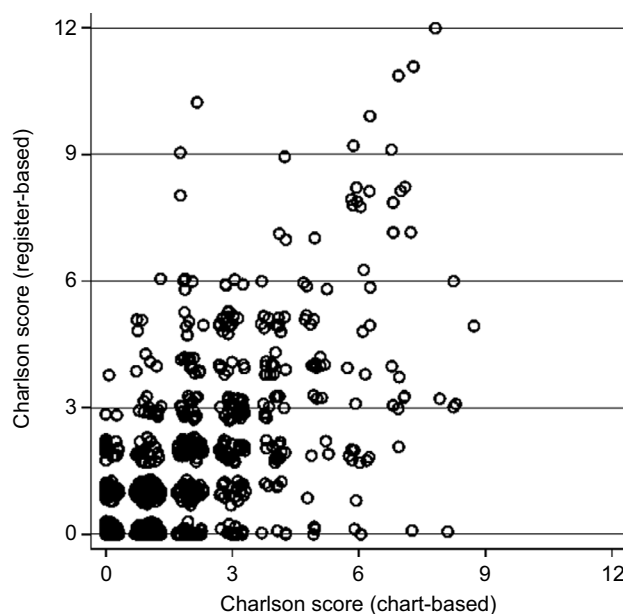

Figure 1 Scatter plot of CCI scores according to source of deriving the score, either from register or chart review.
Note: n=959.
Abbreviation: CCI, Charlson comorbidity index.

Table 3 Cross-tabulation of categorized Charlson index from the 2 sources

Categories from charts	From administrative data			Sum
	0	1	≥2	
0	308	36	20	364
1	88	73	40	201
≥2	60	58	276	394
Sum	456	167	336	959

Notes: McNemar's test, $p<0.001$. Data represent number of patients.

Table 2 Spearman's rank correlations between different variations of Charlson comorbidity scores (n=959)

Correlations	CCI chart (0–33 scale)	CCI chart, 3 categories (0, 1, ≥2)	CCI administrative data, 3 categories (0, 1, ≥2)
Charlson index administrative data (0–29 scale)	0.692	0.668	0.977
Charlson index administrative data, 3 categories (0, 1, ≥2)	0.671	0.662	
Charlson index chart, 3 categories (0, 1, ≥2)	0.966		

Note: $p<0.001$ for all correlations.
Abbreviation: CCI, Charlson comorbidity index.

Table 4 Prevalence, agreement, and positive predictive value for Charlson score from administrative and chart data (n=959)

Condition	International Classification of Diseases, 10th revision codes	Prevalence, n (%)		Agreement, %	Kappa (95% CI)	Positive predictive value (95% CI)#
		Administrative data	Chart			
Myocardial infarction	I21, I22, I25.2	140 (15)	171 (18)	90.7	0.66 (0.59–0.72)	79.3 (71.6–85.7)
Congestive heart failure	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5–I42.9, I43, I50, P29.0	79 (8)	96 (10)	90.9	0.45 (0.36–0.55)	55.7 (44.1–66.9)
Peripheral vascular disease	I70, I71, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9	90 (9)	106 (11)	90.2	0.47 (0.38–0.56)	56.7 (45.8–67.1)
Cerebrovascular disease	G45, G46, I60–I69, H34.0	97 (10)	79 (8)	92.7	0.56 (0.47–0.65)	54.6 (44.2–64.8)
Dementia	F00–F03, G30, F05.1, G31.1	8 (1)	20 (2)	98.3	0.42 (0.19–0.65)	75.0 (34.9–96.8)
Chronic pulmonary disease	I27.8, I27.9, J40–J47, J60–J67, J68.4, J70.1, J70.3	171 (18)	214 (22)	88.8	0.65 (0.59–0.71)	81.3 (74.6–86.8)
Rheumatologic disease	M05, M06, M31.5, M32–M34, M35.1, M35.3, M36.0	12 (1)	28 (3)	97.1	0.29 (0.10–0.47)	50.0 (21.1–78.9)
Peptic ulcer	K25–K28	42 (4)	46 (5)	94.0	0.31 (0.18–0.44)	35.7 (21.6–52.0)
Hemiplegia/paraplegia	G04.1, G11.4, G80.1, G80.2, G81, G82, G83.0, G83.1, G83.2, G83.3, G83.4, G83.9	10 (1)	27 (3)	97.4	0.31 (0.12–0.51)	60.0 (26.2–87.8)
Diabetes without complications	E10.0, E10.1, E10.6, E10.8, E10.9, E11.0, E11.1, E11.6, E11.8, E11.9, E12.0, E12.1, E12.6, E12.8, E12.9, E13.0, E13.1, E13.6, E13.8, E13.9, E14.0, E14.1, E14.6, E14.8, E14.9	101 (11)	92 (10)	92.8	0.60 (0.52–0.69)	61.4 (51.2–70.9)
Diabetes with chronic complications	E10.2–E10.5, E10.7, E11.2–E11.5, E11.7, E12.2–E12.5, E12.7, E13.2–E13.5, E13.7, E14.2–E14.5, E14.7	30 (3)	34 (4)	96.2	0.42 (0.26–0.57)	46.7 (28.3–65.7)
Mild liver disease	B18, K70.0–K70.3, K70.9, K71.3–K71.5, K71.7, K73, K74, K76.0, K76.2–K76.4, K76.8, K76.9, Z94.4	37 (4)	31 (3)	96.0	0.42 (0.27–0.57)	40.5 (24.8–57.9)
Moderate/severe liver disease	I85.0, I85.9, I86.4, I98.2, K70.4, K71.1, K72.1, K72.9, K76.5–K76.7	6 (1)	9 (1)	98.9	0.26 (-0.04–0.56)	33.3 (4.3–77.7)
Renal disease	I12.0, I13.1, N03.2–N03.7, N05.2–N05.7, N18, N19, N25.0, Z49.0–Z49.2, Z94.0, Z99.2	57 (6)	53 (6)	95.6	0.60 (0.48–0.71)	59.6 (45.8–72.4)
Any malignancy (tumor, leukemia, lymphoma)	C00–C26, C30–C34, C37–C41, C43, C45–C58, C60–C76, C81–C85, C88, C90–C97	144 (15)	79 (8)	89.5	0.49 (0.41–0.58)	42.4 (34.2–50.9)
Metastatic solid tumor	C77–C80	16 (2)	30 (3)	97.9	0.56 (0.38–0.73)	81.3 (54.4–96.0)
HIV/AIDS	B20–B22, B24	0 (0)	4 (0.4)	99.6	-	-

Note: #Using chart review as the reference.
Abbreviations: AIDS, acquired immune deficiency syndrome; CI, confidence interval; HIV, human immunodeficiency virus.

value obtained from combining the CCI with SAPS II in the prediction models.

The current study found that more patients had a CCI score >0 when the score was derived from chart review than from the administrative database, which aligns with the findings of previous studies.[12,27] The weighted kappa statistic for agreement between administrative data and chart review for the CCI scores in the current study ranged from 0.65 to 0.67. Previous studies have compared individual comorbidities between administrative data and charts and have reported

Table 5 Discrimination for the risk-adjustment models according to included variables, using *c*-statistics (area under the ROC curve) with 95% CI for 30-day and 1-year mortality

Variables/models	30-day mortality	1-year mortality
Crude mortality, n (%)	249 (26.0)	341 (35.6)
Chart review[a]		
Charlson score	0.656 (0.621–0.692)	0.681 (0.649–0.714)
Charlson score, age, sex	0.743 (0.710 – 0.775)	0.772 (0.743–0.802)
Charlson score, age, sex, type of admission	0.755 (0.723–0.788)	0.783 (0.753–0.812)
Charlson score, age, sex, type of admission, ventilator	0.773 (0.742–0.804)	0.792 (0.763–0.820)
Register-based[b]		
Charlson score	0.607 (0.570–0.644)	0.664 (0.630–0.697)
Charlson score, age, sex	0.734 (0.701–0.768)	0.764 (0.734–0.794)
Charlson score, age, sex, type of admission	0.743 (0.709–0.777)	0.775 (0.745–0.805)
Charlson score, age, sex, type of admission, ventilator	0.762 (0.729–0.794)	0.783 (0.754–0.813)
Physiology-based[a]		
SAPS II	0.821 (0.792–0.850)	0.790 (0.761–0.819)
SAPS II, Charlson score	0.834 (0.806–0.861)	0.816 (0.789–0.843)
SAPS II, Charlson score, age	0.849 (0.822–0.875)	0.838 (0.813–0.864)

Notes: Charlson score categorized as 0, 1, ≥2. All patients (n=959). [a]Charlson score from chart review. [b]Including replacement of 0 for missing Charlson score from register.
Abbreviations: CI, confidence interval; ROC, receiver-operator characteristic; SAPS, Simplified Acute Physiology Score.

Table 6 Model calibration for various combinations of covariates

Model	30-day mortality	*p*-value	1-year mortality	*p*-value
	Hosmer–Lemeshow goodness of fit		Hosmer–Lemeshow goodness of fit	
Chart review				
Charlson score	n/a		n/a	
Charlson score, age, sex	14.79	0.063	11.12	0.195
Charlson score, age, sex, type of admission	9.69	0.287	14.44	0.071
Charlson score, age, sex, type of admission, ventilator	7.42	0.492	4.22	0.836
Register-based[a]				
Charlson score	n/a		n/a	
Charlson score, age, sex	10.42	0.236	9.13	0.331
Charlson score, age, sex, type of admission	4.76	0.783	9.54	0.299
Charlson score 3, age, sex, type of admission, ventilator	2.45	0.964	4.48	0.812
Physiology-based				
SAPS II	5.81	0.669	7.36	0.498
SAPS II, Charlson score (chart review)	18.01	0.021	10.22	0.250
SAPS II, Charlson (chart review), age	10.94	0.205	14.99	0.059

Notes: Charlson score categorized as 0, 1, ≥2 (n=959). n/a, not applicable because only 3 categories are present. [a]Including replacement of 0 for missing Charlson score from register.
Abbreviation: SAPS, Simplified Acute Physiology Score.

simple kappa values of 0.52–0.83[28] and 0.30–0.82,[29] which are comparable with the findings for individual comorbidities in the current study.

The predictive capacity of the CCI from chart review in this study was somewhat better than that of the CCI derived from administrative data, in line with previous studies. For example, comorbidity data from a clinical database of cardiac surgery patients were better at predicting hospital mortality than the data from an administrative database,[30] and the CCI extracted from medical records was better than the CCI derived from administrative billing data for predicting

hospital mortality for Medicare beneficiaries undergoing carotid endarterectomy.[27]

SAPS II scores better predicted 30-day and 1-year mortality than the CCI, which supports previous studies of critical care populations, eg, that physiology-based scores (APACHE II, APACHE III, and SAPS II) have better capacity for predicting hospital mortality[3,16,31] as well as 30-day to 1-year mortality[17] than the CCI. However, models using independently weighted Elixhauser comorbidities discriminated better than an APACHE-weighted score in ICU patients.[32] The finding that the CCI combined with other readily available

variables performed almost as well as physiology-based scores is in concert with previous findings.[17] Moreover, the CCI can be calculated at the time of admission to the ICU, in contrast to physiology-based scores that are dependent on laboratory and bedside clinical data.[18] The *c*-statistic for the CCI was higher for 1-year than for 30-day mortality, as was shown previously in ICU patients.[18]

In the current study, adding comorbidity to SAPS II only marginally improved the ability to predict mortality, supporting previous research indicating that the CCI did not significantly improve the discrimination of the APACHE II weighted score, when it replaced the more rudimentary comorbidity component of the APACHE II score.[15]

There are, however, several notable differences between this study and the studies of comorbidity indexes. Most previous studies are from the US and thus these scored comorbidity based on ICD-9-CM codes, whereas the current study was based on ICD-10 codes, as in Australian and Swiss studies.[11,14] Some studies included diagnosis codes from the index hospital stay or devised algorithms to exclude the main diagnosis or complications.[11–13] This approach is common in studies that use the Elixhauser comorbidity index.[33]

In contrast, the current study only included diagnosis from prior hospital stays or consultations, which is common practice with the CCI; though this is not always clear from the studies. When only using data from prior hospitalizations in 5 selected patient populations, there was little difference in discrimination between the Elixhauser and Charlson/Deyo methods.[34]

Moreover, some studies included patients over long time periods, eg, those admitted to an ICU for >16 years, during which the average age, severity of illness, and comorbidity at the time of admission increased significantly.[35] In contrast, the subjects in the current study were admitted during a 3-year period.

This study was retrospective, as it used chart review and administrative databases, and thus it has inherent limitations with regard to the completeness and quality of the data. Only 1 rater abstracted the presence of the medical conditions from each medical chart, and we did not assess the interrater reliability of the scores abstracted.[21] Medical charts typically include cumulative information on previous diseases and comorbidity; however, in the current study, there is still a possibility that this reference standard was not complete, as we only reviewed charts from 1 hospital. In the Norwegian National Health Service, acute care patients are admitted to a hospital based on their home address, and mobility is fairly low. Moreover, we had administrative data on hospital admissions that went back to 1999, which may explain why

the proportion of patients with previously recorded disease episodes was fairly high. In the current study, the Charlson score from the charts was scored on a 0–33 scale, in contrast to a 0–29 scale used in the algorithm for administrative data. This difference in maximum score of 4 is entirely caused by solid tumor, leukemia, and lymphomas having separate scores of 2 (ie, a maximum of 6 points with all 3 types of malignancy) in the original Charlson model, while in the Deyo/Quan model, these conditions are combined into 1 category for malignancy, which is assigned a weight of 2. In the current study, only 1 patient had 2 of these malignancies (in addition to another comorbidity) and would have been classified as Charlson ≥2 of the categorized Charlson index with both scoring models. Hence, this did not influence the results.

Secondary codes from administrative data do not indicate whether the conditions were preexisting or appeared after an admission; hence, conditions based on laboratory tests may have a higher prevalence in administrative data than in medical charts.[13]

An advantage of using administrative health data for estimating CCI is that it can easily be collected for a large number of subjects, in contrast to more resource-consuming chart reviews. It also enables researchers to conduct and improve risk adjustment in studies where administrative data are available in a register or through linkage with such data, even if access to medical charts is not possible.

The CCI may be incorporated in different ways in statistical analyses. It is most often used as a grouped score, although some authors see this as a limitation and recommend using the continuous score in prediction models,[36] or using the individual comorbidities as dummy variables.[23] Some authors have suggested revising the selection of conditions or the weights in order to improve the predictive capacity of the CCI.[37]

As previously noted, in general, administrative data are accurately coded for severe or life-threatening conditions, but coding may be less accurate for nonspecific or symptomatic conditions.[23] Therefore, results from ICU populations may not be generalizable to other patient populations. Physiology data and physiology-based models are available to a lesser extent for such non-ICU populations, and hence the CCI score may be more useful for prognostic models in these populations.

Conclusion

This study showed that the CCI from chart review resulted in a higher CCI score than that derived from administrative records, although the agreement between the indexes was good. The CCI from both sources predicted 30-day and 1-year mortality in ICU patients. For CCI combined with age, sex,

and type of admission, the capacity to predict 30-day and 1-year mortality approached the level of the physiology-based SAPS II score. The use of the CCI combined with easily collectible variables, such as age, sex, and type of admission, gives useful prognostic information and may complement or replace a physiology-based score in some settings. There seems to be little additional value from combining the CCI with SAPS II in models predicting 30-day or 1-year mortality in ICU patients.

Author contributions

KS made substantial contributions to the conception and design of the study, conducted the analysis and interpreted the data, drafted the paper, and critically revised the manuscript. SAS and HH made substantial contributions to the conception and design of the study, the acquisition of data, and critically commented on the manuscript for important intellectual content. RH made substantial contributions to the study design and data collection and critically commented on the manuscript for scientific content. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Knaus W, Wagner D, Draper E. APACHE III study design: analytic plan for evaluation of severity and outcome in intensive care unit patients. Implications. *Crit Care Med*. 1989;17(12 pt 2):S219–S221.
2. Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993;270(24):2957–2963.
3. Poses RM, McClish DK, Smith WR, Bekes C, Scott WE. Prediction of survival of critically ill patients by admission comorbidity. *J Clin Epidemiol*. 1996;49(7):743–747.
4. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity: a critical review of available methods. *J Clin Epidemiol*. 2003;56(3):221–229.
5. 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.
6. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. *Med Care*. 2004;42(4):355–360.
7. Sharabiani MT, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care*. 2012;50(12):1109–1118.
8. Austin SR, Wong YN, Uzzo RG, Beck JR, Egleston BL. Why summary comorbidity measures such as the Charlson comorbidity index and Elixhauser score work. *Med Care*. 2015;53(9):e65–e72.
9. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992;45(6):613–619.
10. 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.
11. 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.
12. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care*. 2005;43(2):182–188.
13. Chong WF, Ding YY, Heng BH. A comparison of comorbidities obtained from hospital administrative data and medical charts in older patients with pneumonia. *BMC Health Serv Res*. 2011;11:105.
14. Luthi JC, Troillet N, Eisenring MC, et al. Administrative data outperformed single-day chart review for comorbidity measure. *Int J Qual Health Care*. 2007;19(4):225–231.
15. Ho KM, Finn J, Knuiman M, Webb SA. Combining multiple comorbidities with Acute Physiology Score to predict hospital mortality of critically ill patients: a linked data cohort study. *Anaesthesia*. 2007;62(11):1095–1100.
16. Quach S, Hennessy DA, Faris P, Fong A, Quan H, Doig C. A comparison between the APACHE II and Charlson Index Score for predicting hospital mortality in critically ill patients. *BMC Health Serv Res*. 2009;9:129.
17. Christensen S, Johansen MB, Christiansen CF, Jensen R, Lemeshow S. Comparison of Charlson comorbidity index with SAPS and APACHE scores for prediction of mortality following intensive care. *Clin Epidemiol*. 2011;3:203–211.
18. Ladha KS, Zhao K, Quraishi SA, et al. The Deyo-Charlson and Elixhauser-van Walraven Comorbidity Indices as predictors of mortality in critically ill patients. *BMJ Open*. 2015;5(9):e008990.
19. Needham DM, Scales DC, Laupacis A, Pronovost PJ. A systematic review of the Charlson comorbidity index using Canadian administrative databases: a perspective on risk adjustment in critical care research. *J Crit Care*. 2005;20(1):12–19.
20. Hoel H, Skjaker SA, Haagensen R, Stavem K. Decisions to withhold or withdraw life-sustaining treatment in a Norwegian intensive care unit. *Acta Anaesthesiol Scand*. 2014;58(3):329–336.
21. Hall SF, Groome PA, Streiner DL, Rochon PA. Interrater reliability of measurements of comorbid illness should be reported. *J Clin Epidemiol*. 2006;59(9):926–933.
22. Ghali WA, Hall RE, Rosen AK, Ash AS, Moskowitz MA. Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data. *J Clin Epidemiol*. 1996;49(3):273–278.
23. 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(6):676–682.
24. Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*. 1991;100(6):1619–1636.
25. Altman DG. *Practical Statistics for Medical Research*. 1st ed. London: Chapman & Hall; 1991.
26. Hosmer DW Jr, Lemeshow S. *Applied Logistic Regression*. New York, NY: John Wiley & Sons, Inc; 1989.
27. Kieszak SM, Flanders WD, Kosinski AS, Shipp CC, Karp H. A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. *J Clin Epidemiol*. 1999;52(2):137–142.
28. Quan H, Li B, Saunders LD, et al. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. *Health Serv Res*. 2008;43(4):1424–1441.
29. Humphries KH, Rankin JM, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research: are clinical data derived from administrative databases a reliable alternative to chart review? *J Clin Epidemiol*. 2000;53(4):343–349.
30. Hannan EL, Kilburn H Jr, Lindsey ML, Lewis R. Clinical versus administrative data bases for CABG surgery. Does it matter? *Med Care*. 1992;30(10):892–907.
31. Norena M, Wong H, Thompson WD, Keenan SP, Dodek PM. Adjustment of intensive care unit outcomes for severity of illness and comorbidity scores. *J Crit Care*. 2006;21(2):142–150.

32. Johnston JA, Wagner DP, Timmons S, Welsh D, Tsevat J, Render ML. Impact of different measures of comorbid disease on predicted mortality of intensive care unit patients. *Med Care*. 2002;40(10):929–940.

33. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8–27.

34. Stukenborg GJ, Wagner DP, Connors AF Jr. Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. *Med Care*. 2001;39(7):727–739.

35. Williams TA, Ho KM, Dobb GJ, Finn JC, Knuiman MW, Webb SA. Changes in case-mix and outcomes of critically ill patients in an Australian tertiary intensive care unit. *Anaesth Intensive Care*. 2010;38(4):703–709.

36. Hall WH, Ramachandran R, Narayan S, Jani AB, Vijayakumar S. An electronic application for rapidly calculating Charlson comorbidity score. *BMC Cancer*. 2004;4:94.

37. Martins M, Blais R. Evaluation of comorbidity indices for inpatient mortality prediction models. *J Clin Epidemiol*. 2006;59(7):665–669.