Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data

Narayan Sharma (narayan.sharma@unibas.ch)
Instute of Nursing Science, University of Basel https://orcid.org/0000-0001-8698-5469
René Schwendimann
Patient Safety Office, University Hospital Basel, Basel
Olga Endrich
Directorate of Medicine, Inselspital University Hospital Bern, Bern
Dietmar Ausserhofer
College of Health-Care Professions Claudiana
Michael Simon
Institute of Nursing Science (INS), Department of Public Health(DPH), University of Basel

Research article

Keywords: comorbidity indices; weights; risk adjustment; inpatient data; in-hospital mortality

DOI: https://doi.org/10.21203/rs.3.rs-39719/v1

Abstract

Background When chronic conditions are associated with outcomes such as mortality, comorbidity measures are essential both to describe patient health status and to adjust for potential confounding. The Charlson and Elixhauser comorbidity indices are well-established for risk adjustment and mortality prediction. Still, as optimal comorbidity weightings remain undetermined. The present study aimed to derive a set of new population-based Elixhauser comorbidity weightings, then to validate and compare their mortality predictivity against those of the Charlson and Elixhauser-based van Walraven weightings estimates in a population-based cohort.

Methods Retrospective analysis was conducted with routine Swiss general hospital (102 hospitals) data (2012–2017) for 6.09 million inpatient cases. To derive the population-based weightings for the Elixhauser comorbidity index, we randomly halved the inpatient data and validated the results for Part 1 alongside the established weighting systems used for Part 2. Charlson and van Walraven weightings were applied to Charlson and Elixhauser comorbidity indices. Generalized additive models were weighted and adjusted for age, gender and hospital types.

Results Overall, the population-based weights' c-statistic (0.867, 95% CI: 0.865–0.868) was consistently higher than Elixhauser-van Walraven's (0.863, 95% CI: 0.862–0.864) and Charlson's (0.850, 95% CI: 0.849–0.851) in the derivation and validation groups and net reclassification improvement of new weights offers improved predictive performance of 0.4% on the Elixhauser-van Walraven and 6.1% on the Charlson weightings.

Conclusions All weightings were validated with the national dataset and the new population-based weightings model improved the prediction of in-hospital mortality. The newly derive weights support patient population-based analysis of health outcomes.

Background

Critical health outcomes often require effective risk adjustment based on patient characteristics. This is especially true for comorbidities [1, 2], which function as major predictors of mortality [3]. Over one-third of hospitalized patients have at least one comorbidity; two-thirds of those over 65 [4, 2] and three-quarters of those over 85 have at least two [5]. In addition to mortality, comorbidities are associated with lower health-related quality of life, increased disability and higher utilization of both health care services and prescribed medications [6–8].

Data on comorbidities are valuable both for comparison between patient populations and for risk adjustment regarding associated outcomes, especially mortality [9]. Two of the best-known measures are the Charlson Comorbidity Index and the Elixhauser Comorbidity Index [10, 11]. When the Charlson Comorbidity Index was developed in 1987 it included 19 chronic conditions to predict one-year mortality, but has since been shortened to 17. The Elixhauser Comorbidity Index, which was developed in 1998, works on a similar system but includes 30 – or, for some variants, 31 – comorbidities. In addition to in-hospital mortality, it is also designed to predict length of stay and hospital charges [12]. Despite this additional versatility and strong evidence that the Elixhauser Comorbidity Index is statistically superior to the Charlson Comorbidity Index for predicting various

outcomes [13, 14]. Charlson Comorbidity Index continues to be used because of the fewer chronic conditions [15–17] and comparative ease of use in routine situations where time is limited.

Both indices work either via simple (unweighted) sum scores or as weighted scores assigning a risk weight to each comorbidity [6, 18, 19]. A weighted score provides an attractive advantage over plain dummy variables [20, 21], as it reduces the overfitting risk of more parameters, unjustifiable in small datasets [22] and limits computational requirements in large ones [21]. Additionally, evidence indicates that a weighted variable better reduces type I errors than dummy variables while addressing multicollinearity concerns in regression analysis and organizing multiple highly correlated variables into more meaningful information [23, 21]. The weight assigned to each comorbidity reflects a higher, lower or neutral risk of mortality [24]. For example, for hospitalized patients, metastatic cancer entails a considerably higher risk of death than obesity.

To add to the value of early versions of the Elixhauser comorbidities, van Walraven et al. [25] used roughly 13 years' inpatient admission data from one Canadian hospital (1996–2008) to develop a set of weights (VW weights, i.e., the regression coefficient divided by the coefficient in the model with the smallest absolute value) for the 30 Elixhauser comorbidities associated with in-hospital mortality. Using the backward selection and an alpha inclusion criterion of 0.05 to identify independently associated comorbidities, van Walraven identified 21 comorbidities significantly associated with mortality. A VW weight was assigned to each of the 21 Elixhauser comorbidities. Ultimately, VW weights ranged from − 7 to 12, with a weight of 0 assigned to the 9 non-significant comorbidities.

Since then, primarily in North America, studies have used VW weights to predict in-hospital mortality, especially clearly defined patient groups such as surgical, orthopaedic patients, or cancer survivors and those in single sites such as single hospital or intensive care unit (ICU) [25, 26, 21, 14, 12]. Few studies have applied comorbidity adjustments to national or regional inpatient datasets [21, 27]. Therefore, an analysis of a large heterogeneous patient population from a national dataset (Switzerland) is justified both to provide an overview of Elixhauser comorbidities in a European sample and potentially to optimize the comorbidity weights. In addition to increasing the generalizability of these comorbidity weights, the use of such a dataset, representing all hospital inpatients from a very large, heterogeneous patient population, would allow a very accurate comparison of weighting systems. Therefore, the aims of our study were 1) to derive a new population-based comorbidity weighting on a Swiss national inpatient dataset; 2) to validate Charlson, Elixhauser-van Walraven and new weightings on a Swiss national inpatient dataset; and 3) to compare the predictive performance in-hospital mortality of the three weighting systems.

Methods

Study design and population

This is a retrospective population-based analysis of six years' data (2012–2017) from the Swiss national inpatient dataset. Upon our application, subject to a data protection contract (as stipulated by article 22 of the Swiss Federal Act on Data Protection), the Swiss Federal Statistics Office (FSO) provided anonymized data from all Swiss hospital inpatients hospitalized between 2012 and 2017. This included not only general hospitals but also special care (e.g., paediatric, gynaecological) facilities [28]. The FSO classifies general

hospitals (University hospitals, Tertiary hospitals, and three Basic hospitals) into five different levels, based on the number of cases treated per year and/or a special hospital score assigned by Swiss Medical Association ("FMH-Kategorien"). For this study, special care hospitals and children were excluded because of the low levels of comorbidities and the relatively low risk of dying in the hospital [25]. For data protection reasons, age was grouped in five-year groups, and all patients below 20 years of age were excluded. The flowchart for the final adult population included 102 general hospitals (6,094,672 inpatient cases) for the analysis is reported in supplementary figure F1 (Additional file 1).

Dataset and classification of comorbidities indices

The dataset included patient characteristics including sex, age, hospital types, primary and secondary diagnoses based on International Classification of Diseases-10 (ICD-10) codes and hospital discharge information including in-hospital mortality. As condition coding in Switzerland is based on the ICD-10 German Modification (ICD-10 GM), reported in supplementary table S1 (Additional file 1), we used this to identify both Charlson and Elixhauser comorbidities. Specifically, we used Quan et al.'s ICD-10 codes [19] to determine each of the 17 Charlson and 31 Elixhauser comorbidities via the "Comorbidity" package in R [18]. This transforms ICD-10 codes into binary data the relevant comorbidities, their (unweighted) sum scores, and their Charlson and VW-weighted scores.

Descriptive analysis

The study population's general characteristics (sex, age group, hospital types) were classified based on their percentages in the alive and mortality cohorts. The distributions of Charlson and Elixhauser comorbidities, unweighted and weighted scores were computed as percentages of index values of 0, 1–2, and ≥ 3 and < 0, 0, 1–4, and ≥ 5; and as the Charlson weight do not use negative weightings, its weights were calculated for index values of 0, 1–4, and ≥ 5. For each characteristic, standardized mean differences (SMD) between the alive and mortality cohort were provided with SMDs greater than 0.1 considered relevant [29].

Derivation of population-based comorbidity weights

The study population was randomly split into a derivation (50%) and a validation (50%) group. The derivation group was used to determine the adjusted association of all 31 Elixhauser comorbidities with death, treating the anonymous hospital identifier as a random effect [30]. Because of the large sample size, we fitted generalized additive regression models (GAM) to compute the odds ratios (OR) using the package "mgcv" [31] and R programming language, version 3.5.2 [32]. To identify Elixhauser comorbidities associated with in-hospital mortality, we retained variables based on an alpha inclusion criterion of 0.01.

To derive the population-based weightings from the regression model's parameter estimates, we used the method described by Sullivan et al. [33]. Comorbidities not significantly associated with mortality were assigned a weight of zero. The number of (weighted) points assigned to each comorbidity equalled its regression coefficient divided by the coefficient in the model with the smallest absolute value [25, 21, 13, 33] rounded to the nearest whole number. Each person's new Elixhauser comorbidity weighting score was then calculated by summing up all points of all their coded comorbidities.

Validation and comparison of weighted comorbidity models

To validate and compare the performance of the three comorbidity weighting systems, we first created four multivariate in-hospital mortality prediction GAMs for the derivation group. The first model, 'base', contained no comorbidity data – only age group, sex, and hospital types. The other three models used the same variables as the base model, with the first, 'Charlson', using Charlson weightings, the second, 'van Walraven', using the Elixhauser index with van Walraven weightings, and the third, 'population-based weights', using our newly-developed weightings. We then validated all weightings in validation groups by splitting the validation group into six groups by year of discharge. Altogether, then, 24 models (including base models) were created to validate the Charlson, van Walraven, and population-based weights models. An additional four models were created using all cases (combining derivation and validation groups) to evaluate the performance of each model in the total patient population.

We assessed the various comorbidity weightings according to the model performance criteria. Discrimination, i.e., each model's ability to distinguish patients discharged alive from those who died in hospital, was compared using the concordance (c) statistic. The c-statistic quantified each model's ability to assign high probabilities of mortality to patients who died [34]. It's possible values range from 0.50 to 1.0, with 0.50 indicating no ability to discriminate, values less than 0.70 are considered poor, those between 0.70 and 0.80 acceptable, and those of 0.80 or above excellent [35]. Using bootstrap methods, we computed 95% confidence intervals for each c-statistic. We also graphed receiver-operating characteristic curves (ROC) for the visual presentation of the derivation group's c statistics. We compared the three weighted comorbidity models using net reclassification improvement (NRI) for binary models [36–38] from the "nricens" package in R [39] using the Swiss derivation sample. NRI measures the degree of improvement in predicted inpatient mortality probabilities when comorbidity weights are added to the base model [21, 40]. Higher NRI values indicate more accurate reclassification.

Code validation and sensitivity analyses

We also evaluated the R comorbidity package's code handling accuracy in the Swiss setting. To do so we sampled 100 cases and manually reviewed the Swiss ICD-10 codes of the raw data, checking whether the "comorbidity" package had assigned each to the appropriate Charlson or Elixhauser comorbidity. We also performed sensitivity analyses to explore Switzerland's Major Diagnostic Categories' (MDCs') associations, if any, regarding the predictability of in-hospital mortality in combination with the above models and to test whether the combined models' patterns differed from those of uncombined ones. MDCs are 24 mutually exclusive categories into which all primary diagnoses are assigned based on the Swiss diagnostic-related group (DRG) system for hospital reimbursement [41].

Results

Population characteristics

Overall, the adult inpatient population between 2012 and 2017 in all Swiss general hospitals (102) consisted of 6,094,672 cases. The study population characteristics of the adult patients are presented in Table 1. Patients had between 0 and 9 Charlson comorbidities (median 0, interquartile range (IQR): 0–1) and between 0 and 16 Elixhauser comorbidities (median 1, IQR: 0–2). The category of three comorbidity weightings is presented in

supplementary table S2 (Additional file 1). The largest overall SMD for all characteristics (1.322) applied to the Elixhauser index using population-based weights.

Table 1
General characteristics of the total study population

Parameters	Alive cohort (%)	Mortality cohort (%)	SMD
Total population: N = 6,094,672	5,952,005 (97.7)	142,667 (2.3)	
Female	3,280,823 (55.1)	63,912 (44.8)	0.208
Age group			1.006
20–24 years	215,672 (3.6)	292 (0.2)	
25–29 years	327,562 (5.5)	375 (0.3)	
30–34 years	415,022 (7.0)	526 (0.4)	
35–39 years	348,591 (5.9)	718 (0.5)	
40–44 years	299,985 (5.0)	1,368 (1.0)	
45–49 years	350,899 (5.9)	2,503 (1.8)	
50–54 years	408,028 (6.9)	4,312 (3.0)	
55–59 years	430,721 (7.2)	6,503 (4.6)	
60–64 years	466,543 (7.8)	9,068 (6.4)	
65–69 years	528,374 (8.9)	13,322 (9.3)	
70–74 years	554,612 (9.3)	16,899 (11.8)	
75–79 years	535,543 (9.0)	19,888 (13.9)	
80–84 years	509,225 (8.6)	24,853 (17.4)	
85–89 years	365,924 (6.1)	24,042 (16.9)	
90–94 years	161,236 (2.7)	14,156 (9.9)	
95 + years	34,068 (0.6)	3,842 (2.7)	
Hospital types			0.157
University (level 1)	1,078,612 (18.1)	29,379 (20.6)	
Tertiary care (level 2)	3,274,382 (55.0)	83,686 (58.7)	
Basic care (level 3)	736,465 (12.4)	14,863 (10.4)	
Basic care (level 4)	671,182 (11.3)	10,695 (7.5)	
Basic care (level 5)	191,364 (3.2)	4,044 (2.8)	
Number of Charlson comorbidities			1.234
0	3,642,650 (61.2)	17,465 (12.2)	

Abbreviations: SMD standardized mean difference between alive and mortality cohort

Parameters	Alive cohort (%)	Mortality cohort (%)	SMD
1–2	1,907,761 (32.1)	80,876 (56.7)	
>= 3	401,594 (6.7)	44,326 (31.1)	
Number of Elixhauser comorbidities			1.039
0	2,509,169 (42.2)	11,036 (7.7)	
1–2	2,106,780 (35.4)	43,494 (30.5)	
>= 3	1,336,056 (22.4)	88,137 (61.8)	
Abbreviations: SMD standardized mean difference between alive and mortality cohort			

Prevalence of Charlson and Elixhauser comorbidity indices

The most common Charlson comorbidity was any malignancy (including lymphoma and leukaemia, except malignant neoplasm of the skin) in both cohorts, alive (10.2%) and mortality (37.6%). The prevalence for each Charlson comorbidity in the derivation group is presented in supplementary table S3 (Additional file 1).

The most common Elixhauser comorbidities included uncomplicated hypertension (22.7%) in the alive cohort, whereas in the mortality cohort solid tumour without metastasis (33.7%). The prevalence for each Elixhauser comorbidity from the derivation group is presented in Table 2.

Derivation of population-based weights

In the derivation group, two of the 31 Elixhauser comorbidities showed no association with hospital mortality and were removed, leaving 29 in the final model. Sixteen were associated with increased mortality risk, with the strongest associations coming from metastatic cancer (OR: 4.09, 95% CI: 3.98–4.21) and liver disease (OR: 3.83, 95% CI: 3.70–3.97). At the other end of the spectrum, thirteen comorbidities were associated with a decreased risk of hospital mortality. The strongest of these were deficiency anaemia (OR: 0.54, 95% CI: 0.51–0.56) and obesity (OR: 0.59, 95% CI: 0.56–0.63). The adjusted coefficients were used to derive population-based weights with a new maximum weight of 17, for metastatic cancer, and a new minimum of -7, for deficiency anaemia (Table 2).

Table 2
Prevalence, adjusted odds ratio and weights from the (new) population-based derivation sample and the van Walraven (VW) derivation sample [25]

Elixhauser comorbidities	Alive cohort (%)	Mortality cohort (%)	SMD	Adjusted odds ratio (95% CI)		Weights	
	Population-based derivation sample			VW[a]	New[b]	VW[a]	New[b]
Derivation group	2,975,887 (97.7)	71,449 (2.3)					
Congestive heart failure	163,685 (5.5)	16,333 (22.9)	0.514	1.96 (1.85–2.07)	3.07 (3.00–3.14)	7	13
Cardiac arrhythmias	341,280 (11.5)	20,754 (29.0)	0.448	1.71 (1.62–1.80)	1.69 (1.66–1.73)	5	6
Valvular disease	117,450 (3.9)	6,568 (9.2)	0.213	0.91 (0.82–0.99)	0.92 (0.89–0.95)	-1	-1
Pulmonary circulation disorders	53,292 (1.8)	4,813 (6.7)	0.247	1.48 (1.34–1.62)	1.62 (1.57–1.68)	4	6
Peripheral vascular disorders	141,051 (4.7)	6,912 (9.7)	0.192	1.26 (1.17–1.36)	1.27 (1.24–1.31)	2	3
Hypertension (uncomplicated)	676,609 (22.7)	15,692 (22.0)	0.019	–	0.69 (0.68–0.70)	0	-4
Hypertension (complicated)	218,656 (7.3)	11,003 (15.4)	0.256	–	0.79 (0.77–0.81)	0	-3
Paralysis	61,546 (2.1)	5,153 (7.2)	0.246	1.93 (1.75–2.12)	2.60 (2.52–2.69)	7	11
Other neurological disorders	120,045 (4.0)	8,011 (11.2)	0.273	1.83 (1.70–1.96)	2.45 (2.39–2.52)	6	10
Chronic pulmonary disease	170,770 (5.7)	8,269 (11.6)	0.209	1.36 (1.29–1.44)	1.31 (1.27–1.34)	3	3

Abbreviations: SMD standardized mean difference between alive and mortality cohort, *New*[b] population-based, *VW*[a] van Walraven, – excluded in final model

Note: Total cohort % exceed 100% for each cohort, as comorbidities are mutually inclusive

Elixhauser comorbidities	Alive cohort (%)	Mortality cohort (%)	SMD	Adjusted odds ratio (95% CI)		Weights	
Diabetes, uncomplicated	245,817 (8.3)	9,059 (12.7)	0.145	–	1.09 (1.06–1.11)	0	1
Diabetes, complicated	66,161 (2.2)	2,763 (3.9)	0.096	–	0.89 (0.86–0.93)	0	-1
Hypothyroidism	126,062 (4.2)	3,454 (4.8)	0.029	–	0.76 (0.74–0.79)	0	-3
Renal failure	289,047 (9.7)	20,526 (28.7)	0.497	1.63 (1.54–1.73)	2.06 (2.02–2.11)	5	8
Liver disease	49,916 (1.7)	5,822 (8.1)	0.303	2.97 (2.73–3.22)	3.83 (3.7–3.97)	11	16
Peptic ulcer disease, excluding bleeding	5,808 (0.2)	258 (0.4)	0.032	–	–	0	0
AIDS/HIV	2,300 (0.1)	85 (0.1)	0.013	–	–	0	0
Lymphoma	25,049 (0.8)	1,759 (2.5)	0.127	2.55 (2.31–2.81)	2.19 (2.07–2.31)	9	9
Metastatic cancer	119,667 (4.0)	18,907 (26.5)	0.657	3.30 (3.10–3.52)	4.09 (3.98–4.21)	12	17
Solid tumour without metastasis	268,298 (9.0)	24,046 (33.7)	0.631	1.47 (1.39–1.56)	2.36 (2.3–2.42)	4	10
Rheumatoid arthritis/collagen vascular diseases	47,305 (1.6)	1,254 (1.8)	0.013	–	0.91 (0.86–0.97)	0	-1
Coagulopathy	90,551 (3.0)	9,528 (13.3)	0.382	1.30 (1.22–1.40)	2.12 (2.07–2.18)	3	9
Obesity	68,155 (2.3)	1,011 (1.4)	0.065	0.64 (0.53–0.77)	0.59 (0.56–0.63)	-4	-6

Abbreviations: SMD standardized mean difference between alive and mortality cohort, *New*[b] population-based, *VW*[a] van Walraven, — excluded in final model

Note: Total cohort % exceed 100% for each cohort, as comorbidities are mutually inclusive

Elixhauser comorbidities	Alive cohort (%)	Mortality cohort (%)	SMD	Adjusted odds ratio (95% CI)		Weights	
Weight loss	98,545 (3.3)	9,527 (13.3)	0.369	1.85 (1.67–2.04)	1.67 (1.63–1.71)	6	6
Fluid and electrolyte disorders	257,618 (8.7)	17,440 (24.4)	0.434	1.61 (1.53–1.69)	1.58 (1.55–1.61)	5	5
Blood loss anaemia	19,759 (0.7)	685 (1.0)	0.033	0.81 (0.70–0.93)	0.66 (0.60–0.71)	-2	-5
Deficiency anaemia	72,290 (2.4)	1,886 (2.6)	0.013	0.80 (0.71–0.90)	0.54 (0.51–0.56)	-2	-7
Alcohol abuse	96,708 (3.2)	3,086 (4.3)	0.056	–	0.75 (0.72–0.78)	0	-3
Drug abuse	38,044 (1.3)	583 (0.8)	0.045	0.50 (0.42–0.60)	0.67 (0.61–0.73)	-7	-5
Psychoses	29,598 (1.0)	404 (0.6)	0.049	–	0.72 (0.65–0.79)	0	-4
Depression	173,898 (5.8)	3,715 (5.2)	0.028	0.73 (0.67–0.80)	0.73 (0.70–0.75)	-3	-3

Abbreviations: SMD standardized mean difference between alive and mortality cohort, *New*[b] population-based, *VW*[a] van Walraven, — excluded in final model

Note: Total cohort % exceed 100% for each cohort, as comorbidities are mutually inclusive

Validation and comparison of weighted comorbidity models

All three comorbidity weighting systems (Charlson, Elixhauser van Walraven and population-based) indicated higher in-hospital mortality risk than the base model. They also showed equivalent discrimination in the derivation and validation groups (Table 3). Overall, the Charlson weighting model's c-statistic was 0.850 (95% CI: 0.849–0.851), the van Walraven model was 0.863 (95% CI: 0.862–0.864) and population-based weights model was 0.867 (95% CI: 0.865–0.868), with similar results for both the derivation and validation groups. The population-based weights model's discrimination was statistically superior with some c-statistic variability across the six years' data. The comorbidity model using the population-based weights showed discrimination power almost 10% higher than the base models across the derivation, validation, and all case groups.

Table 3
Performance measures of base, Charlson, van Walraven and population-based weights models for in-hospital mortality in derivation, validation and all cases groups

C Statistic (95% CI)								
	Derivation group	Validation groups						All cases
	n = 3,047,336	n_1 = 491,962	n_2 = 496,684	n_3 = 504,741	n_4 = 514,267	n_5 = 520,277	n_6 = 519,405	N = 6,094,672
	All years (2012–2017)	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	All years (2012–2017)
Base model	0.757 (0.755–0.759)	0.758 (0.782–0.789)	0.758 (0.754–0.762)	0.756 (0.753–0.760)	0.752 (0.748–0.756)	0.756 (0.752–0.759)	0.750 (0.746–0.754)	0.757 (0.755–0.759)
Charlson weights model	0.850 (0.847–0.851)	0.849 (0.846–0.852)	0.852 (0.849–0.855)	0.854 (0.851–0.857)	0.849 (0.846–0.852)	0.854 (0.851–0.857)	0.844 (0.841–0.847)	0.850 (0.849–0.851)
VW weights model	0.863 (0.862–0.864)	0.862 (0.859–0.865)	0.866 (0.863–0.869)	0.867 (0.864–0.869)	0.863 (0.860–0.866)	0.869 (0.866–0.872)	0.862 (0.859–0.864)	0.863 (0.862–0.864)
Population-based weights model	0.867 (0.865–0.868)	0.865 (0.862–0.868)	0.869 (0.866–0.871)	0.871 (0.868–0.873)	0.866 (0.863–0.869)	0.872 (0.869–0.874)	0.865 (0.862–0.867)	0.867 (0.865–0.868)
Abbreviations: VW van Walraven, *CI* confidence interval								

Base model: age groups, sex, hospital types

Charlson weights model: base and Charlson weights

VW weights model: base and Elixhauser/ van Walraven weights

Population-based weights model: base and Elixhauser/ Population-based weights

As shown in Fig. 1, the population-based weights model's discrimination was substantially better than the Charlson's or base model's, and slightly better than the van Walraven's. The NRI figures confirmed this (Table 4). Compared with the base model, that using population-based weights increased predicted probabilities of mortality by 31.0% and decreased the predicted probabilities of alive by 25.5%, resulting in a higher NRI (0.369, 95% CI: 0.363–0.369) than either the Charlson or the Elixhauser van Walraven model. Additionally, a sensitivity analysis using MDCs did not offer any improvements in models' performance than without MDCs.

Table 4
Net Reclassification Improvement (NRI) compared to base model in derivation group

Derivation group					
	NRI (95% CI)	Mortality increased Pr(Up|Case)	Alive increased Pr(Up|Ctrl)	Mortality decreased Pr(Down|Case)	Alive decreased Pr(Down|Ctrl)
Base model	Reference	Reference	Reference	Reference	Reference
Charlson weights model	0.308 (0.301–0.308)	0.255 (0.250–0.255)	0.071 (0.071–0.072)	0.145 (0.145–0.148)	0.269 (0.269–0.270)
VW weights model	0.365 (0.360–0.367)	0.299 (0.294– 0.300)	0.079 (0.079–0.080)	0.115 (0.114–0.117)	0.262 (0.261–0.262)
Population-based weights model	0.369 (0.363–0.369)	0.310 (0.305–0.310)	0.087 (0.086–0.087)	0.108 (0.107–0.110)	0.255 (0.254–0.255)
Abbreviations: NRI Net Reclassification Improvement, *CI* confidence interval					

Pr(Up, Down) | (Case, Ctrl) represent the proportion of patients whose predicted probabilities increased or decreased for in-hospital mortality and alive cohorts respectively

NRI= (Pr(Up|Case) - Pr(Down|Case)) + (Pr(Down|Ctrl) - Pr(Up|Ctrl))

Base model: age group, sex, hospital types

Charlson weights model: base and Charlson weights

VW weights model: base and Elixhauser/ van Walraven weights

Population-based weights model: base and Elixhauser/ Population-based weights

Discussion

This study used a six-year dataset on a Swiss multi-million-patient population to explore Charlson and Elixhauser comorbidities' capacities to predict in-hospital mortality. We first derived a set of optimal population-based weightings for the 31 Elixhauser comorbidities using the national inpatient dataset. Our population-based comorbidity weights were validated – along with the other two weighting systems in the validation sample of Swiss national inpatient data – indicated in-hospital mortality risks with greater sensitivity than either of these weighting systems. The new weighting set had a robust association with in-hospital mortality and discriminated equally well in the derivation and validation groups.

Comparing predictivity regarding in-hospital mortality, the optimized population-based weightings performed slightly better than the Charlson and Elixhauser-van Walraven sets. However, it also supplied weights for eight Elixhauser comorbidities (e.g. diabetes, hypertension, and psychosis) eliminated by van Walraven et al. (2009) [25]. Of the risk-associated comorbidities retained in both the van Walraven and the population-based weights, several comorbidities showed similar results, e.g., the highest odds ratios to metastatic cancer and liver disease. And regarding the comorbidities with negative associations, only slight differences were observed between the van Walraven and population-based weights (e.g., hypothyroidism or obesity were likely to be healthier, and have better survival).

From an epidemiological perspective, chronic diseases such as cancer, heart and liver disease increase the risk of dying in hospitals and certain other disorders, e.g., anaemia and hypothyroidism, actually reduce that risk. These results are in line with those of Zellweger et al.'s [42] study using the Swiss national death registry of hospital inpatient data from 2010–2012. Furthermore, van Walraven et al.'s [25] via a single Canadian hospital's records and Thompson et al., [21] using Maryland State inpatient data, showed similar results. These relations could insight the global burden of in-hospital mortality is due to the rising chronic diseases.

The existing weighting systems [11, 25, 21, 14] represent data from a specific geographical region, patient group, or even limited numbers of hospitals or settings, matching the generalizability of these weighting systems remained difficult. As this study addresses such issues, with a large dataset representing the Swiss inpatient population, it provides population-based comorbidity adjustments for the prediction of mortality or other health outcomes. The slightly improved performance of the population-based weights system suggests that it might be worthwhile to derive country- or region-specific comorbidity weights from representative patient populations.

C-statistics and ROCs are widely used to assess predictive performance. Nonetheless, one downside of comparing c-statistic and ROCs is that differences between c-statistics are often small, [43] as it was the case when we compared our new weights and van Walraven's. Over the past decade, it has become common to use NRIs to compare different models' performance, even though it might differ with the cut-offs taken for analysis [44, 37]. In our study, taking the same cut-offs, NRI calculations confirmed the three weighting systems' rankings i.e., population-based, van Walraven and Charlson weightings.

The primary strength of this study was the large sample size and the heterogeneity of the Swiss inpatient population across all general hospitals over six years, which made it representative of the entire country. To our knowledge, this study is the first to derive and validate Elixhauser weightings in Swiss hospital inpatient data. We used standard regression methodology for large datasets, including random effects at the hospital level, and internally validated our models. We also used accepted methods to modify our adjusted model into a population-based weightings system that re-includes the association of several comorbidities formerly excluded from the Elixhauser index [33]. Despite differences in individual comorbidities' prevalence and weightings, Charlson, Elixhauser/VW, and the population-based weights performed well across the derivation, validation, and all-cases groups. We also used NRIs, allowing a robust comparison of model performance. Finally, the methods we applied were explicit and can be replicated by other researchers, who can adjust or control for patient comorbidity via their own and national databases.

Our study also has certain notable limitations. We derived our weights using statistical criteria, while clinical knowledge might be needed to determine each comorbidity's value. Since we used codes assigned in routine data, the capture of the comorbidities could be influenced by other factors, such as physician and nurse documentation, code assignment accuracy, and the possibility that capture of comorbidities is biased towards those for which the Swiss DRG / MDC pays more [45, 41]. Additionally, as Swiss data protection regulations prevented us from obtaining the inpatients' exact age, we could not differentiate children exactly under 18 years and could not specify each year. This also might have influenced the predictive accuracy of the tested models.

Conclusions

We found that Charlson, Elixhauser/van Walraven and population-based weights all indicated very similar risks of in-hospital mortality. However, regarding the national inpatient dataset, the population-based weights' results were more robust across derivation and validation groups and showed slightly higher discriminative performance for risk adjustment than either the Charlson or the Elixhauser/van Walraven weightings. In short, the population-based weights allowed improved in-hospital mortality prediction. Researchers wishing to use the Elixhauser comorbidities should utilize the population-based weights derived in this study. Most importantly, given access to similar data, researchers can use the methods described here to derive their own country- or region-specific comorbidity weights.

Abbreviations

ECI	Elixhauser Comorbidity Index
CCI	Charlson Comorbidity Index
ICD-10 GM	International Classification of Diseases version-10 German Modification
DRGs	Diagnosis-Related Groups
SMD	Standardized Mean Differences
NRI	Net Reclassification Improvement
MDCs	Major Diagnostic Categories
ROC	Receiver-Operating Characteristic
FSO	Federal Statistical Office
SBK	Swiss Nurses' Association
GAM	Generalized Additive Model
FMH	Swiss Medical Association
CI	Confidence Interval

Declarations

Ethics approval and consent to participate

Further ethical approval was deemed unnecessary, as the study was subjected to data protection contract (as stipulated by article 22 of the Swiss Federal Act on Data Protection) with Swiss Federal Statistics Office. Consent to participate was not applicable.

Consent for publication

Not applicable.

Availability of data and material

Upon application, the data that support the findings of this study are available from Federal Statistical Office (FSO), Switzerland.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by Swiss Nurses' Association (SBK). The funders had no role in the design, conduct, analysis or reporting of this study.

Authors' contributions

This paper is derived from the PhD dissertation research in the field of health service and patient safety of the first author. NS and MS had full access to all of the data in the study and take responsibility for the integrity of the data and accuracy of the analysis. NS, RS, OE, DA, and MS contributed to the conception and design of the study. NS drafted the manuscript and all the authors contributed substantially to the interpretation, visualization of the data, critically revised, edited the manuscript for important intellectual content and agreed to be accountable for all aspects. All of the authors listed above approved this version of the manuscript to be published.

Acknowledgements

We would like to thank the Swiss Federal Statistics Office (FSO) and its data management team for their support in providing inpatient data for our analyses. We are especially grateful to all the Swiss hospitals for their regular data reporting and information forwarded to the FSO and we would like to thank Mr. Chris Shultis for English editing of this paper. We are also thankful to the Swiss Nurses' Association (SBK) for their financial support for this study.

References

1. Harrison C, Britt H, Miller G, Henderson J. Examining different measures of multimorbidity, using a large prospective cross-sectional study in Australian general practice. BMJ open. 2014;4(7):e004694.
2. Pefoyo AJK, Bronskill SE, Gruneir A, et al. The increasing burden and complexity of multimorbidity. BMC Public Health. 2015;15(1):415.
3. Metcalfe D, Masters J, Delmestri A, et al. Coding algorithms for defining Charlson and Elixhauser co-morbidities in Read-coded databases. BMC medical research methodology. 2019;19(1):115.
4. Wallace E, Salisbury C, Guthrie B, Lewis C, Fahey T, Smith SM. Managing patients with multimorbidity in primary care. Bmj. 2015;350:h176.
5. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012;380(9836):37–43.
6. Fraccaro P, Kontopantelis E, Sperrin M, et al. Predicting mortality from change-over-time in the Charlson Comorbidity Index: A retrospective cohort study in a data-intensive UK health system. Medicine. 2016;95(43).
7. Moore BJ, White S, Washington R, Coenen N, Elixhauser A. Identifying increased risk of readmission and in-hospital mortality using hospital administrative data. Medical care. 2017;55(7):698–705.
8. Divo M, Cote C, de Torres JP, et al. Comorbidities and risk of mortality in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2012;186(2):155–61.
9. Toson B, Harvey LA, Close JC. The ICD-10 Charlson Comorbidity Index predicted mortality but not resource utilization following hip fracture. J Clin Epidemiol. 2015;68(1):44–51.
10. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care. 1998:8–27.
11. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373–83.
12. Chang H-J, Chen P-C, Yang C-C, Su Y-C, Lee C-C. Comparison of Elixhauser and Charlson methods for predicting oral cancer survival. Medicine. 2016;95(7).
13. Fortin Y, Crispo JA, Cohen D, McNair DS, Mattison DR, Krewski D. External validation and comparison of two variants of the Elixhauser comorbidity measures for all-cause mortality. PloS one. 2017;12(3):e0174379.
14. Menendez ME, Neuhaus V, Van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. Clinical Orthopaedics Related Research®. 2014;472(9):2878–86.
15. Dailiana Z, Papakostidou I, Varitimidis S, Michalitsis S, Veloni A, Malizos K. Surgical treatment of hip fractures: factors influencing mortality. Hippokratia. 2013;17(3):252.
16. Shetty KD, Bhattacharya J. Changes in hospital mortality associated with residency work-hour regulations. Ann Intern Med. 2007;147(2):73–80.
17. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. Medical care. 2004;42(4):355–60.

18. Gasparini A. comorbidity: An R package for computing comorbidity scores. Journal of Open Source Software. 2018;3(23):648.
19. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care. 2005:1130–9.
20. Te Grotenhuis M, Pelzer B, Eisinga R, Nieuwenhuis R, Schmidt-Catran A, Konig R. When size matters: advantages of weighted effect coding in observational studies. International Journal of Public Health. 2017;62(1):163–7.
21. Thompson NR, Fan Y, Dalton JE, et al. A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. Medical care. 2015;53(4):374.
22. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events. Bmj. 2015;351:h3868.
23. Song M-K, Lin F-C, Ward SE, Fine JP. Composite variables: when and how. Nursing research. 2013;62(1):45.
24. Pule ML, Buckley E, Niyonsenga T, Roder D. The effects of comorbidity on colorectal cancer mortality in an Australian cancer population. Scientific reports. 2019;9(1):8580.
25. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical care. 2009:626–33.
26. Austin SR, Wong Y-N, Uzzo RG, Beck JR, Egleston BL. Why summary comorbidity measures such as the Charlson comorbidity index and Elixhauser score work. Medical care. 2015;53(9):e65.
27. Stanley J, Sarfati D. The new measuring multimorbidity index predicted mortality better than Charlson and Elixhauser indices among the general population. J Clin Epidemiol. 2017;92:99–110.
28. De CP, Camenzind P, Sturny I, et al. Switzerland: health system review. Health systems in transition. 2015;17(4):1-288, xix.
29. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivar Behav Res. 2011;46(3):399–424.
30. Wilson M, Welch J, Schuur J, O'Laughlin K, Cutler D. Hospital and emergency department factors associated with variations in missed diagnosis and costs for patients age 65 years and older with acute myocardial infarction who present to emergency departments. Acad Emerg Med. 2014;21(10):1101–8.
31. Li Z, Wood SN. Faster model matrix crossproducts for large generalized linear models with discretized covariates. Statistics and Computing. 2019:1–7.
32. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2019.
33. Sullivan LM, Massaro JM, D'Agostino Sr RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Statistics in medicine. 2004;23(10):1631–60.
34. 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):676–82.
35. Pencina MJ, D'Agostino RB. Evaluating discrimination of risk prediction models: the C statistic. Jama. 2015;314(10):1063–4.

36. Pencina MJ, D'Agostino Sr RB, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Statistics in medicine. 2008;27(2):157–72.
37. Pencina MJ, D'Agostino Sr RB, D'Agostino RB Jr, Vasan RS. Comments on 'Integrated discrimination and net reclassification improvements—Practical advice'. Stat Med. 2008;27(2):207–12.
38. Pencina MJ, D'Agostino Sr RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Statistics in medicine. 2011;30(1):11–21.
39. Inoue E, Inoue ME. Net Reclassification Improvement (NRI) for Risk Prediction Models with Time to Event and Binary Response Data. 2018.
40. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol. 2011;64(7):749–59.
41. Wasserfallen J-B, Zufferey J. Financial impact of introducing the Swiss-DRG reimbursement system on potentially avoidable readmissions at a university hospital. Swiss medical weekly. 2015;145.
42. Zellweger U, Junker C, Bopp M, Group SNCS. Cause of death coding in Switzerland: evaluation based on a nationwide individual linkage of mortality and hospital in-patient records. Population health metrics. 2019;17(1):2.
43. Loging WT. Bioinformatics and Computational Biology in Drug Discovery and Development: Cambridge University Press; 2016.
44. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med. 2014;160(2):122–31.
45. Quan H, Eastwood C, Cunningham CT, et al. Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). BMJ open. 2013;3(10):e003716.

Figures

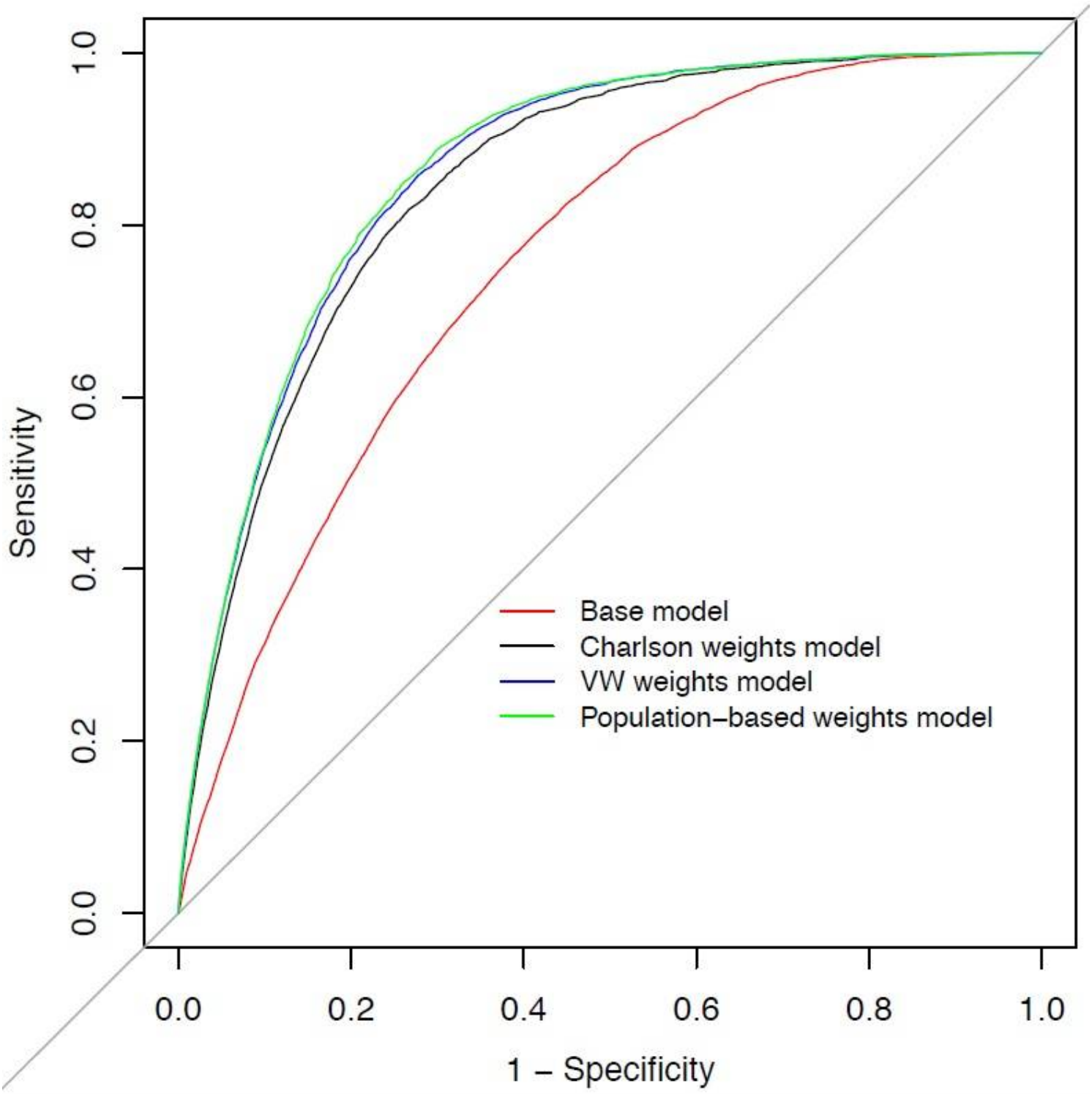

Figure 1

Receiver-operating characteristic (ROC) curves for generalized additive models predicting in-hospital mortality. Base model: age group, sex, hospital types; Charlson weights model: base and Charlson weights; VW weights model: base and Elixhauser/ van Walraven weights; Population-based weights model: base and Elixhauser/ population-based weights

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx