Journal of Medical Systems (2022) 46:37
https://doi.org/10.1007/s10916-022-01805-3

HEALTH POLICY

All Patient Refined-Diagnosis Related Groups' (APR-DRGs) Severity of Illness and Risk of Mortality as predictors of in-hospital mortality

João Vasco Santos[1,2,3,4] · João Viana[1,2] · Carla Pinto[5] · Júlio Souza[1,2] · Fernando Lopes[1,2] · Alberto Freitas[1,2] · Sílvia Lopes[6,7]

Received: 3 November 2021 / Accepted: 7 February 2022

Abstract

The aims of this study were to assess All-Patient Refined Diagnosis-Related Groups' (APR-DRG) Severity of Illness (SOI) and Risk of Mortality (ROM) as predictors of in-hospital mortality, comparing with Charlson Comorbidity Index (CCI) and Elixhauser Comorbidity Index (ECI) scores. We performed a retrospective observational study using mainland Portuguese public hospitalizations of adult patients from 2011 to 2016. Model discrimination (C-statistic/ area under the curve) and goodness-of-fit (R-squared) were calculated. Our results comprised 4,176,142 hospitalizations with 5.9% in-hospital deaths. Compared to the CCI and ECI models, the model considering SOI, age and sex showed a statistically significantly higher discrimination in 49.6% (132 out of 266) of APR-DRGs, while in the model with ROM that happened in 33.5% of APR-DRGs. Between these two models, SOI was the best performer for nearly 20% of APR-DRGs. Some particular APR-DRGs have showed good discrimination (e.g. related to burns, viral meningitis or specific transplants). In conclusion, SOI or ROM, combined with age and sex, perform better than more widely used comorbidity indices. Despite ROM being the only score specifically designed for in-hospital mortality prediction, SOI performed better. These findings can be helpful for hospital or organizational models benchmarking or epidemiological analysis.

Keywords In-hospital mortality · Prediction · Diagnosis-Related Groups · Severity of Illness · Risk of Mortality · Charlson Comorbidity Index · Elixhauser Comorbidity Index

Introduction

Predicting mortality can lead to better patient evaluation, clinical information for differential and tailored interventions, comparison of providers' performance and even allocation of resources [1]. Specifically, in-hospital mortality is typically regarded to meet these ends, thereby being of utmost importance from a clinical, epidemiological or managerial perspective.

Prediction of in-hospital mortality is traditionally more linked to the intensive care context, with several scores or indices being developed and used in that specific field, such as the acute physiology and chronic health evaluation (APACHE), the simplified acute physiology score II (SAPS-II) or the sequential organ failure assessment (SOFA) [2]. In a broader context within inpatient care, scores or indices based on the patients' underlying comorbidities have been developed and widely used for mortality prediction in different groups of diseases or types of care, with the Charlson and Elixhauser Comorbidity Indices or its adaptations being frequently used [1, 3].

Topical Collection on *Health Policy*

✉ João Vasco Santos
jvasco.santos@gmail.com

1 MEDCIDS – Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal

2 CINTESIS – Centre for Health Technology and Services Research, Porto, Portugal

3 Public Health Unit, ACES Grande Porto VIII – Espinho/Gaia, ARS Norte, Vila Nova de Gaia, Portugal

4 NOVA National School of Public Health, Universidade NOVA de Lisboa, Lisboa, Portugal

5 Department of Biomedicine, Faculty of Medicine, University of Porto, Porto, Portugal

6 NOVA National School of Public Health, Public Health Research Centre, Universidade NOVA de Lisboa, Lisboa, Portugal

7 Comprehensive Health Research Center, Lisboa, Portugal

Published online: 06 May 2022

The Charlson Comorbidity Index (CCI) defines 17 weighted comorbidities, whereas the Elixhauser Charlson Index (ECI) includes 30 unweighted comorbidities [4, 5]. Both methods work either through unweighted sum of scores or via weighted scores, with the latter approach assigning a risk weight to each patient's comorbidity [3]. These indices are based on International Classification of Diseases (ICD) coding systems and have been updated several times regarding the list of specific comorbidities, codes and weights [6–11].

Moreover, these measures rely on diagnostic information that is typically available in hospital administrative databases, such as those used for prospective payment systems based upon Diagnosis Related Groups (DRGs). Since the introduction of DRG hospital reimbursement in 1983 in the United States (US), several countries have gradually introduced this system as their main hospital provider payment mechanism [12–15], and several modifications and DRG versions have been developed since then [16]. Any DRG-based hospital payment system, however, includes an exhaustive patient classification algorithm, where the hospital episode is assigned to a clinically and economically homogeneous group (DRG) based upon the diagnoses, procedures and demographic characteristics, such as age and sex. Episodes within the same DRG group are expected to present a similar clinical evolution.

In 1990, the APR (All Patient-Refined) DRG version introduced a new concept of patient stratification to increase the granularity on patient characteristics and, thus, provide a better predictive model for resource use, by assigning a severity of illness (SOI) subclass and a risk of mortality (ROM) subclass to each episode, in addition to the base DRG. Both SOI and ROM are DRG-specific and depend on other patient's underlying characteristics, namely comorbidities [17]. The APR-DRGs were employed for reimbursement purposes in some European countries, namely Portugal, Spain, Belgium and Italy, some Arab countries and in over 30 states in the US [18]. Each APR-DRG is subdivided into four SOI and four ROM subclasses, ranging from 1 to 4 (e.g.1 - minor, 2 - moderate, 3 - major, and 4 - extreme) within each DRG. Both SOI and ROM are calculated separately based mostly on secondary diagnoses and their interaction with age, main diagnosis and selected procedures. SOI determines the overall patient illness severity according to the extent of physiological decomposition or loss of organ system function, while ROM estimates the likelihood of in-hospital mortality [17].

Apart from their role in hospital financing, APR-DRGs' SOI and ROM have been previously used as predictive models for in-hospital mortality among patients with specific diseases (e.g. heart failure or stroke), in specific settings (e.g. intensive care units), among surgical patients or for global hospital admissions [19–24]. However, the two latter scenarios are not suitable for using SOI and ROM as mortality predictors, mainly because SOI and ROM have been developed for risk-stratification within DRG-specific contexts (i.e. discrimination of severity of illness and risk of mortality among patients within the same DRG). In this sense, understanding how these scores contribute to patient mortality within specific clinical contexts is essential to describe their usefulness and relevance for risk adjustment, especially in comparison with other metrics (e.g., CCI or ECI). To the best of our knowledge there are no studies assessing APR-DRGs' SOI and ROM as predictive models for in-hospital mortality specific to a APR-DRG or group of APR-DRGs.

Therefore, the aims of this study were (1) to assess SOI and ROM as predictive models of in-hospital mortality in different clinical contexts (DRGs), (2) to compare SOI and ROM with CCI and ECI scores; and (3) to identify the APR-DRGs in which SOI and/or ROM perform better.

Methods

Study design and data

This study was a retrospective observational nationwide study using inpatient data from Portuguese public hospitalizations, selected from a mainland database provided by the Portuguese Central Administration the Health System (ACSS).

Hospitalizations of adult patients (18 years old or older) with a hospital discharge date between 1st January 2011 and 31st December 2016, with diagnoses and procedures coded in the International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM), were included.

We excluded episodes with transfers to other hospitals [127,766 (2.9%) out of 4,467,199; no APR-DRGs excluded] or "leaving against medical advice" [33,990 (0.8%) out of 4,339,433; 1 APR-DRG excluded out of 285]. APR-DRGs with only 3 in-hospital deaths or less [87,923 (2.0%) out of 4,305,443; 18 APR-DRGs out of 284] were also excluded, as well as episodes with a length of stay equal to 0 [41,378 (1.0%) out of 4,217,520; no APR-DRGs excluded]. All APR-DRGs not considered in this study due to inclusion or exclusion criteria (N=48) are listed in Supplementary Table 1.

The database comprises demographic and clinical data at episode (hospital case) level for hospitals in mainland Portugal. Age, gender, discharge date, discharge status and diagnoses were considered. The APR-DRG, SOI and ROM that which episode was assigned to is also available in the database.

The 3M APR-DRG version 31 software was previously used to compute the SOI and ROM of each hospitalization.

Outcome and models

The two most commonly used generic scores for in-hospital mortality prediction are the Charlson Comorbidity Index (CCI) [4] and the Elixhauser Comorbidity Index (ECI) [5]. The first was developed by Charlson et al. in 1987, defining 17 comorbidities to estimate mortality risk, while the second was developed by Elixhauser in 1998 considering 30 comorbidities [4, 5]. These indices have several versions of comorbidity weights: original [4], Schneeweiss's [10] and Quan's updates [7] for CCI; and original [5], Van Walraven's [8], Moore's [9] and Thompson's updates [11] for ECI. Weights for some comorbidities are negative in the ECI updates [8, 9, 11].

Statistical analysis

We performed descriptive statistics and crossed frequencies between SOI and ROM. Additionally, we analysed the proportion distributions of in-hospital deaths by SOI and ROM level, overall and by APR-DRG.

A total of eleven logistic regression models were computed for each APR-DRG using in-hospital death as the outcome (dependent variable): model 1- Original CCI (+ age + sex); model 2- Schneeweiss' updated CCI (+ age + sex); model 3- Quan's updated CCI (+ age + sex); model 4- Original ECI (+ age + sex); model 5- Van Walraven's updated ECI (+ age + sex); model 6- Moore's updated ECI (+ age + sex); model 7- Thompson's updated ECI (+ age + sex); model 8- SOI; model 9- SOI (+ age + sex); model 10- ROM; and model 11- ROM (+ age + sex).

Models were evaluated for model discrimination and goodness-of-fit. Model discrimination was assessed by plotting the Receiver Operating Characteristic (ROC) curves and by calculating the area under the curve, i.e. c-statistic. This can range between 0.5 (random discrimination) and 1 (perfect discrimination). Bootstrapping with 100 samples and bias-correction were performed, allowing the quantification of uncertainty by returning 95% confidence interval estimates [25]. In fact, for less frequent APR-DRGs, only limited data would be available for assessing models using test sets.

Goodness-of-fit was assessed by calculating the Nagelkerke R-square measure. Bootstrapping with 100 samples was performed in order to obtain 95% confidence interval estimates for Nagelkerke R-square measures.

Descriptive and inferential statistics were performed using Microsoft Excel v15.3, IBM SPSS Statistics v24.0 and R v3.6.2.

Results

From the 4,176,142 hospitalizations among adults in mainland Portugal public hospitals between 2011 and 2016, there were 286,297 in-hospital deaths — 5.9% of all hospitalizations. A total of 2,338,303 (56.0%) hospitalizations were females, while age ranged between 18 and 108 years-old with a median of 64 years-old. The distribution by SOI levels were 2,613,046 (62.6%), 1,000,688 (24.0%), 456,361 (10.9%) and 106,047 (2.5%) for levels 1, 2, 3 and 4, respectively, with associated in-hospital mortality rates of 0.6%, 8.0%, 21.7% and 50.6%. The distribution by ROM levels were 2,038,704 (48.8%), 1,475,760 (35.3%), 567,451 (13.6%) and 94,227 (2.3%) for levels 1, 2, 3 and 4, respectively, with associated in-hospital mortality rates of 0.8%, 6.0%, 18.2% and 43.2%. Table 1 displays the summary descriptive measures for each CCI and ECI.

Model 9 (SOI + age + sex) had the highest discrimination (c-statistic) and goodness-of-fit (R-square) for 75.6% APR-DRGs (201 out of 266), followed by model 11 (ROM + age + sex) with 16.5% (44 out of 266) – Fig. 1. For model 9,

Table 1 Summary descriptive measures for Charlson Comorbidity Index (CCI), Elixhauser Comorbidity Index (ECI) and updates among Portuguese public hospitalizations coded in ICD-9-CM between 2011 and 2016, for the entire sample, those with in-hospital death as outcome and those without in-hospital death as outcome

	Entire sample (N = 4,176,142)	**Episodes without in-hospital death** (N = 3,889,845)	**Episodes with in-hospital death** (N = 286,297)
	min **p25** median **p75** max; mean (SD)	min **p25** median **p75** max; mean (SD)	min **p25** median **p75** max; mean (SD)
Original CCI	0.0 **0.0** 0.0 **2.0** 18.0; 1.1 (1.8)	0.0 **0.0** 0.0 **1.0** 17.0; 1.0 (1.6)	0.0 **1.0** 2.0 **4.0** 18.0; 2.8 (2.5)
Schneeweiss' updated CCI	0.0 **0.0** 0.0 **2.0** 20.0; 1.4 (2.2)	0.0 **0.0** 0.0 **2.0** 20.0; 1.3 (2.0)	0.0 **1.0** 3.0 **6.0** 20.0; 3.4 (2.8)
Quan's updated CCI	0.0 **0.0** 0.0 **1.0** 15.0; 0.9 (1.6)	0.0 **0.0** 0.0 **1.0** 15.0; 0.8 (1.5)	0.0 **0.0** 2.0 **4.0** 15.0; 2.5 (2.5)
Original ECI	0.0 **0.0** 1.0 **2.0** 14.0; 1.5 (1.7)	0.0 **0.0** 1.0 **2.0** 14.0; 1.5 (1.7)	0.0 **1.0** 3.0 **4.0** 14.0; 2.9 (1.9)
Van Walraven's updated ECI	-16.0 **0.0** 0.0 **6.0** 58.0; 3.5 (6.2)	-16.0 **0.0** 0.0 **5.0** 58.0; 3.0 (5.8)	-11.0 **5.0** 10.0 **16.0** 58.0; 10.6 (7.9)
Moore's updated ECI	-22.0 **0.0** 0.0 **5.0** 66.0; 3.0 (7.2)	-22.0 **-1.0** 0.0 **4.0** 63.0; 2.5 (6.6)	-16.0 **3.0** 11.0 **18.0** 66.0; 11.5 (9.9)
Thompson's updated ECI	-27.0 **0.0** 0.0 **8.0** 82.0; 4.3 (9.0)	-27.0 **0.0** 0.0 **7.0** 80.0; 3.6 (6.6)	-19.0 **6.0** 15.0 **18.0** 82.0; 14.9 (11.8)

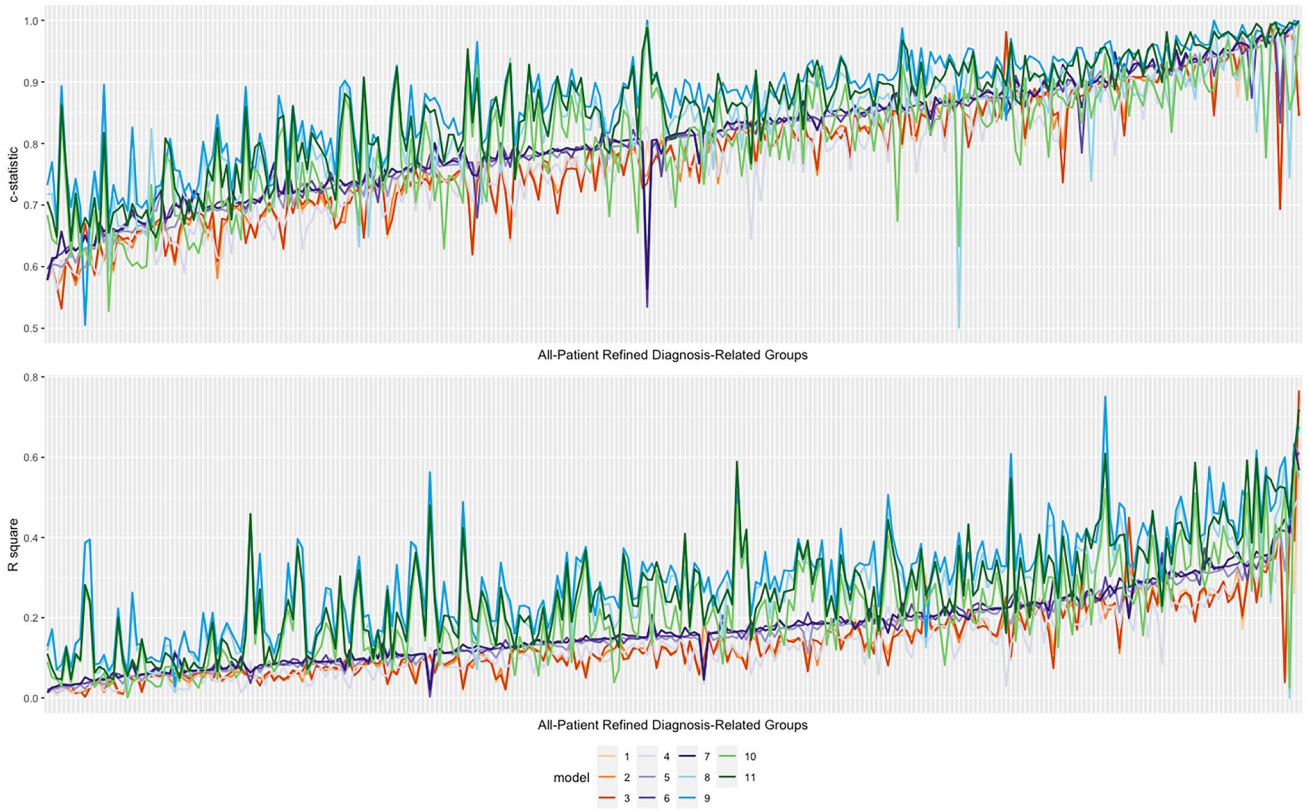

Fig. 1 Model discrimination (c-statistic) and goodness-of-fit (R-square) for each of the 11 models by APR-DRG, ordered by the median value for each APR-DRG, among Portuguese public hospitalizations coded in ICD-9-CM between 2011 and 2016. Models: model 1 - Original Charlson Comorbidity Index, CCI (+ age + sex); model 2 – Schneeweiss' updated CCI (+ age + sex); model 3 - Quan's updated CCI (+ age + sex); model 4 - Original Elixhauser Comorbidity Index, ECI (+ age + sex); model 5 - Van Walraven's updated ECI (+ age + sex); model 6 - Moore's updated ECI (+ age + sex); model 7 - Thompson's updated ECI (+ age + sex); model 8 – Severity of Illness, SOI; model 9 - SOI (+ age + sex); model 10 – Risk of Mortality, ROM; and model 11 - ROM (+ age + sex)

c-statistic ranged between 0.505 and 1.000 (mean=0.872; median=0.891; SD=0.083) and R-square ranged between 0.028 and 0.750 (mean=0.298; median=0.302; SD=0.128). For model 11, c-statistic ranged between 0.639 and 0.997 (mean=0.854; median=0.870; SD=0.090) and R-square ranged between 0.046 and 0.719 (mean=0.261; median=0.255; SD=0.131). C-statistic was higher than 0.900 for 113 (42.5%) and 90 (33.8%) out of 266 APR-DRGs for models 9 and 11, respectively.

Compared to all CCI and ECI models (models 1 to 7), model 9 showed a statistically significantly higher discrimination in 49.6% (132 out of 266) of APR-DRGs and a statistically significantly higher goodness-of-fit in 57.9% (154 out of 266) of APR-DRGs. Moreover, model 11 had a statistically significantly higher discrimination in 33.5% (89 out of 266) of APR-DRGs and a statistically significantly higher goodness-of-fit in 43.6% (116 out of 266) of APR-DRGs, compared to all CCI and ECI models.

Between the two best performing models, model 9 showed a statistically significantly higher discrimination in 21.8% (58 out of 266) APR-DRGs and goodness-of-fit in 22.9% (61 out of 266) APR-DRGs than model 11. In the opposite direction, mode 11 showed a statistically significantly higher discrimination in 2 out of 266 APR-DRGs and goodness-of-fit in 1 out of 266 APR-DRGs compared with model 9.

Tables 2 and 3 present model discrimination (c-statistic) and goodness-of-fit (R-square) respectively, for all models by APR-DRG, including the results for the 5 most frequent APR-DRGs, the 5 APR-DRGs with the highest in-hospital mortality and top-5 and bottom-5 APR-DRGs for each models' performance measure.

Model 9 presented the highest predictive discrimination ability considering the top-5 APR-DRGs with the highest in-hospital mortality, and for the majority of APR-DRGs within the top-5 most frequent, top-5 c-statistic APR-DRGs and bottom-5 c-statistic APR-DRGs (Table 2). Furthermore, for the top-5 APR-DRGs with highest mortality, all models including models 9 and 11presented a good discrimination ability (c-statistic > 0.6), although these values were close to the other models.

Table 2 Model discrimination (c-statistic) for all 11 models by APR-DRG, for the 5 most frequent APR-DRGs, the 5 APR-DRGs with the highest in-hospital mortality and top-5 and bottom-5 APR-DRGs for each measure, among Portuguese public hospitalizations coded in ICD-9-CM between 2011 and 2016. Models: model 1 - Original Charlson Comorbidity Index, CCI (+ age + sex); model 2 – Schneeweiss' updated CCI (+ age + sex); model 3 - Quan's updated CCI (+ age + sex); model 4 - Original Elixhauser Comorbidity Index, ECI (+ age + sex); model 5 - Van Walraven's updated ECI (+ age + sex); model 6 - Moore's updated ECI (+ age + sex); model 7 - Thompson's updated ECI (+ age + sex); model 8 – Severity of Illness, SOI; model 9 - SOI (+ age + sex); model 10 – Risk of Mortality, ROM; and model 11 - ROM (+ age + sex)

APR-DRG	N (in-hospital mortality %)	Model 1 (95%CI)	Model 2 (95%CI)	Model 3 (95%CI)	Model 4 (95%CI)	Model 5 (95%CI)	Model 6 (95%CI)	Model 7 (95%CI)	Model 8 (95%CI)	Model 9 (95%CI)	Model 10 (95%CI)	Model 11 (95%CI)
Top-5 APR-DRGs by maximum c-statistic												
514 - female reproductive system reconstructive procedures	24056 (0.02)	0.873 (0.716;0.996)	0.866 (0.768;0.999)	0.845 (0.614;0.998)	0.963 (0.822;0.995)	0.941 (0.843;0.995)	0.909 (0.763;0.997)	0.946 (0.744;0.999)	0.998 (0.995;1.000)	1.000 (0.999;1.000)	0.888 (0.628;1.000)	0.982 (0.963;1.000)
363 - breast procedures except mastectomy	40310 (0.01)	0.976 (0.960;1.000)	0.974 (0.959;0.999)	0.966 (0.935;0.980)	0.944 (0.901;0.998)	0.988 (0.983;1.000)	0.990 (0.971;1.000)	0.995 (0.985;0.998)	0.998 (0.992;1.000)	0.999 (0.999;1.000)	0.886 (0.683;1.000)	0.993 (0.983;0.999)
540 - cesarean delivery	120469 (0.01)	0.803 (0.618;0.986)	0.748 (0.623;0.987)	0.736 (0.582;0.992)	0.826 (0.626;0.965)	0.595 (0.422;0.948)	0.535 (0.514;0.979)	0.564 (0.492;0.971)	0.999 (0.999;1.000)	0.999 (0.999;1.000)	0.992 (0.979;1.000)	0.988 (0.979;1.000)
561 - postpartum & post abortion diagnoses w/o procedure	120469 (0.01)	0.942 (0.902;0.998)	0.845 (0.612;1.000)	0.845 (0.620;1.000)	0.995 (0.989;1.000)	0.999 (0.995;1.000)	0.999 (0.998;1.000)	0.999 (0.997;1.000)	0.996 (0.985;1.000)	0.997 (0.989;1.000)	0.994 (0.987;0.999)	0.997 (0.991;1.000)
483 - testes & scrotal procedures	8160 (0.23)	0.958 (0.920;0.977)	0.953 (0.918;0.974)	0.959 (0.927;0.977)	0.918 (0.845;0.978)	0.973 (0.951;0.987)	0.963 (0.938;0.985)	0.974 (0.955;0.991)	0.993 (0.989;0.997)	0.997 (0.995;0.999)	0.991 (0.984;0.997)	0.996 (0.992;0.998)
Bottom-5 APR-DRGs by maximum c-statistic												
144 - respiratory signs, symptoms & minor diagnoses	82738 (9.31)	0.649 (0.642;0.656)	0.637 (0.630;0.642)	0.646 (0.640;0.651)	0.625 (0.618;0.631)	0.658 (0.651;0.663)	0.673 (0.668;0.679)	0.670 (0.665;0.676)	0.662 (0.658;0.668)	0.690 (0.685;0.697)	0.613 (0.609;0.619)	0.667 (0.661;0.673)
720 - septicemia & disseminated infections	40130 (35.01)	0.624 (0.620;0.631)	0.620 (0.615;0.626)	0.626 (0.620;0.631)	0.618 (0.612;0.624)	0.632 (0.627;0.637)	0.634 (0.630;0.641)	0.634 (0.628;0.639)	0.654 (0.649;0.658)	0.688 (0.685;0.694)	0.602 (0.597;0.609)	0.665 (0.656;0.668)
6 - pancreas transplant	96 (4.17)	0.621 (0.592;1.000)	0.599 (0.508;1.000)	0.670 (0.625;0.986)	0.590 (0.594;1.000)	0.629 (0.682;1.000)	0.648 (0.633;1.000)	0.651 (0.616;0.995)	0.518 (0.454;0.974)	0.505 (0.631;1.000)	0.630 (0.499;0.995)	0.679 (0.592;1.000)
5 - tracheostomy w mv 96+ hours w/o extensive procedure	2130 (54.55)	0.655 (0.632;0.675)	0.659 (0.638;0.681)	0.654 (0.629;0.677)	0.654 (0.632;0.678)	0.658 (0.634;0.677)	0.662 (0.640;0.683)	0.665 (0.646;0.692)	0.594 (0.569;0.614)	0.699 (0.642;0.690)	0.528 (0.511;0.551)	0.655 (0.635;0.675)
41 - nervous system malignancy	8285 (23.56)	0.569 (0.555;0.582)	0.567 (0.555;0.583)	0.568 (0.555;0.582)	0.563 (0.549;0.579)	0.603 (0.592;0.621)	0.617 (0.604;0.635)	0.615 (0.603;0.633)	0.661 (0.650;0.671)	0.667 (0.651;0.678)	0.637 (0.624;0.651)	0.648 (0.635;0.663)
Top-5 APR-DRGs by frequency												
560 - vaginal delivery	282920 (0.001)	0.872 (0.686;1.000)	0.872 (0.774;1.000)	0.872 (0.686;1.000)	0.881 (0.706;1.000)	0.866 (0.681;1.000)	0.866 (0.680;1.000)	0.866 (0.678;1.000)	0.501 (0.501;0.501)	0.870 (0.679;1.000)	0.633 (0.633;0.634)	0.910 (0.777;1.000)
139 - other pneumonia	200755 (19.04)	0.658 (0.655;0.661)	0.653 (0.650;0.656)	0.657 (0.653;0.659)	0.650 (0.647;0.652)	0.666 (0.663;0.669)	0.676 (0.673;0.678)	0.674 (0.672;0.677)	0.661 (0.657;0.664)	0.702 (0.700;0.704)	0.607 (0.605;0.609)	0.684 (0.681;0.687)
194 - heart failure	122411 (11.55)	0.644 (0.639;0.648)	0.643 (0.639;0.648)	0.645 (0.639;0.649)	0.639 (0.634;0.643)	0.657 (0.654;0.661)	0.666 (0.661;0.670)	0.661 (0.657;0.665)	0.665 (0.661;0.669)	0.711 (0.708;0.715)	0.636 (0.632;0.640)	0.701 (0.697;0.706)
540 - cesarean delivery	120469 (0.01)	0.803 (0.618;0.986)	0.748 (0.623;0.987)	0.736 (0.582;0.992)	0.826 (0.626;0.965)	0.595 (0.422;0.948)	0.535 (0.514;0.979)	0.564 (0.492;0.971)	0.999 (0.999;1.000)	0.999 (0.999;1.000)	0.992 (0.979;1.000)	0.988 (0.979;1.000)

Table 2 (continued)

APR-DRG	N (in-hospital mortality %)	Model 1 (95%CI)	Model 2 (95%CI)	Model 3 (95%CI)	Model 4 (95%CI)	Model 5 (95%CI)	Model 6 (95%CI)	Model 7 (95%CI)	Model 8 (95%CI)	Model 9 (95%CI)	Model 10 (95%CI)	Model 11 (95%CI)
45 - cva & precerebral occlusion w infarct	105082 (11.52)	0.671 (0.667;0.676)	0.674 (0.670;0.678)	0.676 (0.671;0.681)	0.669 (0.664;0.674)	0.690 (0.684;0.694)	0.699 (0.695;0.704)	0.704 (0.700;0.707)	0.772 (0.768;0.776)	0.800 (0.795;0.804)	0.709 (0.705;0.713)	0.765 (0.761;0.768)
Top-5 APR-DRGs by in-hospital mortality												
196 - cardiac arrest	653 (57.58)	0.579 (0.536;0.616)	0.579 (0.539;0.633)	0.578 (0.541;0.617)	0.582 (0.550;0.631)	0.584 (0.547;0.623)	0.596 (0.567;0.640)	0.578 (0.539;0.620)	0.717 (0.686;0.745)	0.732 (0.700;0.776)	0.684 (0.644;0.727)	0.706 (0.664;0.747)
5 - tracheostomy w mv 96+ hours w/o extensive procedure	2130 (54.55)	0.655 (0.632;0.675)	0.659 (0.638;0.681)	0.654 (0.629;0.677)	0.654 (0.632;0.678)	0.658 (0.634;0.677)	0.662 (0.640;0.683)	0.665 (0.646;0.692)	0.594 (0.569;0.614)	0.699 (0.642;0.690)	0.528 (0.511;0.551)	0.655 (0.635;0.675)
4 - tracheostomy w mv 96+ hours w extensive procedure or ecmo	2746 (42.90)	0.656 (0.640;0.675)	0.657 (0.640;0.681)	0.657 (0.634;0.677)	0.664 (0.646;0.685)	0.662 (0.639;0.679)	0.666 (0.646;0.686)	0.669 (0.655;0.690)	0.650 (0.632;0.666)	0.707 (0.691;0.725)	0.601 (0.585;0.615)	0.692 (0.669;0.710)
130 - respiratory system diagnosis w ventilator support 96+ hours	5298 (42.64)	0.651 (0.640;0.668)	0.643 (0.632;0.658)	0.651 (0.639;0.666)	0.637 (0.622;0.652)	0.656 (0.644;0.670)	0.657 (0.645;0.674)	0.659 (0.645;0.674)	0.666 (0.656;0.683)	0.701 (0.688;0.715)	0.639 (0.624;0.649)	0.693 (0.677;0.706)
382 - malignant breast disorders	5257 (37.97)	0.671 (0.657;0.682)	0.670 (0.654;0.682)	0.678 (0.667;0.690)	0.621 (0.606;0.634)	0.718 (0.704;0.731)	0.724 (0.714;0.740)	0.724 (0.713;0.737)	0.729 (0.714;0.740)	0.732 (0.717;0.743)	0.665 (0.651;0.675)	0.676 (0.663;0.692)

Table 3 Model goodness-of-fit (R-square) for all 11 models by APR-DRG, for the 5 most frequent APR-DRGs, the 5 APR-DRGs with the highest in-hospital mortality and top-5 and bottom-5 APR-DRGs for each measure, among Portuguese public hospitalizations coded in ICD-9-CM between 2011 and 2016. Models: model 1 - Original Charlson Comorbidity Index, CCI (+ age + sex); model 2 – Schneeweiss' updated CCI (+ age + sex); model 3 - Quan's updated CCI (+ age + sex); model 4 - Original Elixhauser Comorbidity Index, ECI (+ age + sex); model 5 - Van Walraven's updated ECI (+ age + sex); model 6 - Moore's updated ECI (+ age + sex); model 7 - Thompson's updated ECI (+ age + sex); model 8 – Severity of Illness, SOI; model 9 - SOI (+ age + sex); model 10 – Risk of Mortality, ROM; and model 11 - ROM (+ age + sex)

APR-DRG	N (in-hospital mortality %)	Model 1 (95%CI)	Model 2 (95%CI)	Model 3 (95%CI)	Model 4 (95%CI)	Model 5 (95%CI)	Model 6 (95%CI)	Model 7 (95%CI)	Model 8 (95%CI)	Model 9 (95%CI)	Model 10 (95%CI)	Model 11 (95%CI)
Top-5 APR-DRGs by maximum R square												
160 - major cardiothoracic repair of heart anomaly	92 (4.35)	0.605 (0.386;1.000)	0.750 (0.571;1.000)	0.766 (0.579;1.000)	0.501 (0.328;1.000)	0.608 (0.344;1.000)	0.611 (0.410;1.000)	0.567 (0.272;1.000)	0.545 (0.146;1.000)	0.677 (0.417;1.000)	0.552 (0.227;0.816)	0.719 (0.507;1.000)
514 - female reproductive system reconstructive procedures	24056 (0.02)	0.220 (0.054;0.549)	0.198 (0.024;0.435)	0.231 (0.040;0.544)	0.229 (0.135;0.424)	0.246 (0.108;0.511)	0.301 (0.073;0.791)	0.265 (0.111;0.588)	0.707 (0.380;0.977)	0.750 (0.542;1.000)	0.521 (0.146;.860)	0.609 (0.317;1.000)
561- postpartum & post abortion diagnoses w/o procedure	3788 (0.11)	0.261 (0.090;0.548)	0.449 (0.007;0.943)	0.391 (0.013;0.851)	0.476 (0.250;0.814)	0.635 (0.420;1.000)	0.595 (0.425;1.000)	0.631 (0.455;1.000)	0.603 (0.215;0.862)	0.626 (0.207;1.000)	0.591 (0.375;0.786)	0.604 (0.419;1.000)
483 - testes & scrotal p rocedures	8160 (0.23)	0.238 (0.182;0.313)	0.244 (0.182;0.330)	0.238 (0.197;0.317)	0.288 (0.168;0.398)	0.351 (0.260;0.468)	0.321 (0.237;0.460)	0.365 (0.271;0.461)	0.582 (0.480;0.720)	0.617 (0.525;0.745)	0.551 (0.433;0.670)	0.597 (0.495;0.742)
651 - other procedures of blood & blood-forming organs	2413 (0.83)	0.094 (0.049;0.189)	0.111 (0.068;0.206)	0.103 (0.039;0.213)	0.134 (0.096;0.250)	0.184 (0.106;0.337)	0.250 (0.109;0.409)	0.228 (0.106;0.395)	0.605 (0.412;0.753)	0.608 (0.506;0.757)	0.534 (0.330;0.709)	0.548 (0.445;0.751)
Bottom-5 APR-DRGs by maximum R square												
893 - hiv w multiple significant hiv related conditions	664 (13.86)	0.020 (0.008;0.058)	0.019 (0.010;0.058)	0.019 (0.007;0.054)	0.020 (0.007;0.052)	0.028 (0.012;0.072)	0.035 (0.013;0.076)	0.033 (0.017;0.072)	0.066 (0.030;0.117)	0.083 (0.051;0.133)	0.031 (0.015;0.076)	0.046 (0.024;0.098)
422 - hypovolemia & related electrolyte disorders	11588 (11.93)	0.027 (0.020;0.035)	0.017 (0.014;0.023)	0.031 (0.025;0.038)	0.011 (0.008;0.015)	0.029 (0.023;0.034)	0.032 (0.027;0.039)	0.032 (0.026;0.039)	0.071 (0.061;0.081)	0.077 (0.068;0.091)	0.036 (0.029;0.044)	0.047 (0.039;0.057)
41 - nervous system malignancy	8285 (23.56)	0.010 (0.007;0.014)	0.010 (0.006;0.013)	0.010 (0.006;0.015)	0.008 (0.005;0.013)	0.022 (0.018;0.029)	0.030 (0.022;0.035)	0.028 (0.024;0.037)	0.070 (0.059;0.082)	0.071 (0.059;0.083)	0.046 (0.039;0.055)	0.049 (0.041;0.058)
144 - respiratory signs, symptoms & minor diagnoses	82738 (9.31)	0.040 (0.036;0.043)	0.033 (0.030;0.036)	0.039 (0.036;0.042)	0.027 (0.025;0.030)	0.045 (0.042;0.049)	0.053 (0.049;0.056)	0.051 (0.048;0.055)	0.055 (0.051;0.059)	0.068 (0.064;0.071)	0.029 (0.026;0.032)	0.052 (0.048;0.056)
5 - tracheostomy w mv 96+ hours w/o extensive procedure	2130 (54.55)	0.055 (0.037;0.075)	0.059 (0.047;0.079)	0.054 (0.039;0.070)	0.054 (0.042;0.071)	0.058 (0.046;0.086)	0.060 (0.048;0.077)	0.063 (0.048;0.080)	0.021 (0.012;0.032)	0.066 (0.048;0.084)	0.001 (0.000;0.006)	0.055 (0.038;0.073)
Top-5 APR-DRGs by frequency												
560 - vaginal delivery	282920 (0.001)	0.434 (0.020;1.000)	0.434 (0.012;1.000)	0.434 (0.018;1.000)	0.437 (0.035;1.000)	0.417 (0.019;1.000)	0.415 (0.021;0.903)	0.412 (0.011;1.000)	0.0002 (0.0002;0.0002)	0.433 (0.017;1.000)	0.026 (0.023;0.027)	0.451 (0.051;1.000)
139 - other pneumonia	200755 (19.04)	0.055 (0.053;0.056)	0.050 (0.049;0.052)	0.054 (0.052;0.056)	0.048 (0.047;0.050)	0.060 (0.058;0.061)	0.066 (0.064;0.067)	0.065 (0.062;0.067)	0.065 (0.063;0.067)	0.090 (0.087;0.092)	0.029 (0.028;0.030)	0.073 (0.071;0.075)

Table 3 (continued)

APR-DRG	N (in-hospital mortality %)	Model 1 (95%CI)	Model 2 (95%CI)	Model 3 (95%CI)	Model 4 (95%CI)	Model 5 (95%CI)	Model 6 (95%CI)	Model 7 (95%CI)	Model 8 (95%CI)	Model 9 (95%CI)	Model 10 (95%CI)	Model 11 (95%CI)
194 - heart failure	122411 (11.55)	0.036 (0.034;0.039)	0.036 (0.034;0.039)	0.037 (0.035;0.039)	0.033 (0.032;0.036)	0.044 (0.042;0.047)	0.049 (0.047;0.052)	0.047 (0.044;0.049)	0.064 (0.061;0.067)	0.086 (0.081;0.088)	0.044 (0.040;0.047)	0.078 (0.075;0.082)
540 - cesarean delivery	120469 (0.01)	0.111 (0.026;0.277)	0.107 (0.014;0.209)	0.107 (0.015;0.281)	0.125 (0.055;0.329)	0.053 (0.009;0.190)	0.003 (0.003;0.118)	0.022 (0.004;0.139)	0.561 (0.451;0.662)	0.563 (0.473;0.668)	0.478 (0.317;0.649)	0.480 (0.333;0.712)
45 - cva & precerebral occlusion w infarct	105082 (11.52)	0.053 (0.050;0.056)	0.055 (0.052;0.058)	0.056 (0.053;0.059)	0.052 (0.049;0.055)	0.066 (0.062;0.069)	0.071 (0.068;0.074)	0.075 (0.072;0.078)	0.177 (0.172;0.182)	0.187 (0.182;0.193)	0.104 (0.099;0.109)	0.135 (0.130;0.141)
Top-5 APR-DRGs by in-hospital mortality												
196 - cardiac arrest	653 (57.58)	0.014 (0.006;0.038)	0.014 (0.006;0.031)	0.014 (0.005;0.036)	0.016 (0.005;0.035)	0.015 (0.007;0.044)	0.021 (0.009;0.047)	0.014 (0.004;0.037)	0.118 (0.080;0.163)	0.128 (0.085;0.167)	0.089 (0.060;0.139)	0.110 (0.082;0.160)
5 - tracheostomy w mv 96+ hours w/o extensive procedure	2130 (54.55)	0.055 (0.037;0.075)	0.059 (0.047;0.079)	0.054 (0.039;0.070)	0.054 (0.042;0.071)	0.058 (0.046;0.086)	0.060 (0.048;0.077)	0.063 (0.048;0.080)	0.021 (0.012;0.032)	0.066 (0.048;0.084)	0.001 (0.000;0.006)	0.055 (0.038;0.073)
4 - tracheostomy w mv 96+ hours w extensive procedure or ecmo	2746 (42.90)	0.053 (0.043;0.069)	0.054 (0.040;0.065)	0.053 (0.036;0.067)	0.059 (0.047;0.073)	0.059 (0.049;0.075)	0.063 (0.051;0.077)	0.066 (0.055;0.081)	0.063 (0.048;0.076)	0.102 (0.085;0.120)	0.039 (0.028;0.052)	0.090 (0.066;0.105)
130 - respiratory system diagnosis w ventilator support 96+ hours	5298 (42.64)	0.055 (0.047;0.069)	0.049 (0.041;0.060)	0.055 (0.045;0.064)	0.044 (0.036;0.053)	0.058 (0.049;0.068)	0.058 (0.051;0.070)	0.060 (0.050;0.069)	0.067 (0.058;0.081)	0.096 (0.083;0.113)	0.046 (0.037;0.055)	0.088 (0.079;0.100)
382 - malignant breast disorders	5257 (37.97)	0.087 (0.073;0.100)	0.083 (0.070;0.091)	0.093 (0.081;0.110)	0.024 (0.018;0.031)	0.112 (0.099;0.124)	0.119 (0.107;0.136)	0.119 (0.108;0.134)	0.137 (0.124;0.151)	0.138 (0.127;0.154)	0.077 (0.063;0.086)	0.080 (0.068;0.094)

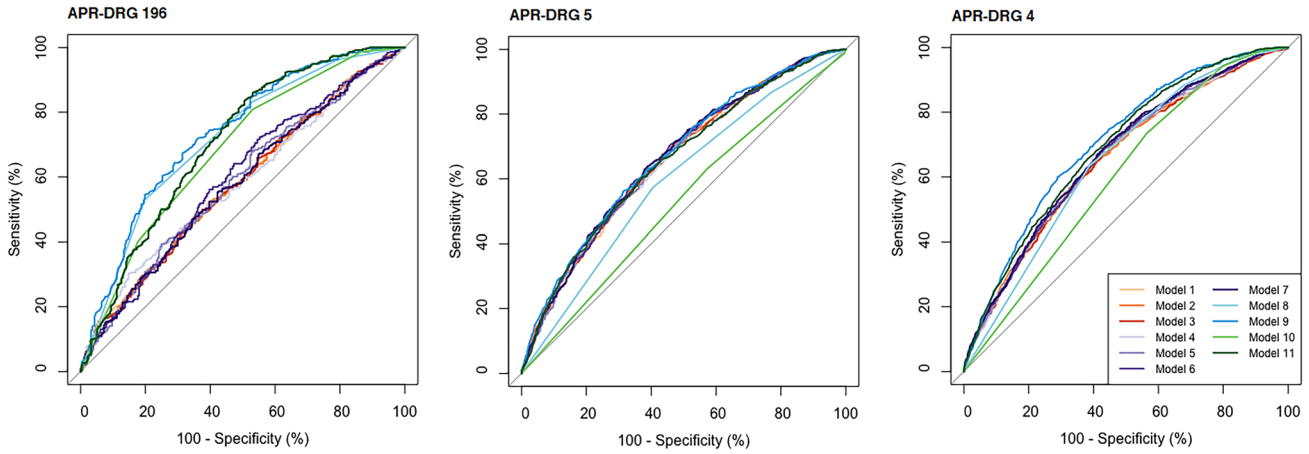

Fig. 2 ROC curves for all 11 models by APR-DRG for the 3 with the highest in-hospital mortality rate (i.e. APR-DRG 196, 5 and 4), among Portuguese public hospitalizations coded in ICD-9-CM between 2011 and 2016. Models: model 1 - Original Charlson Comorbidity Index, CCI (+ age + sex); model 2 – Schneeweiss' updated CCI (+ age + sex); model 3 - Quan's updated CCI (+ age + sex); model 4 - Original Elixhauser Comorbidity Index, ECI (+ age + sex); model 5 - Van Walraven's updated ECI (+ age + sex); model 6 - Moore's updated ECI (+ age + sex); model 7 - Thompson's updated ECI (+ age + sex); model 8 – Severity of Illness, SOI; model 9 - SOI (+ age + sex); model 10 – Risk of Mortality, ROM; and model 11 - ROM (+ age + sex)

The amount of variation of in-hospital mortality explained by the several scores, as indicated by the R-square, was poor for the top-5 most frequent and top-5 APR-DRGs with the highest in-hospital mortality (Table 3), where models explained less than 10% of variability in death for most APR-DRGs within these rankings. The highest R-square (0.766 [0.579-1.000]) was obtained with model 3 for the APR-DRG 160 – "major cardiothoracic repair of heart anomaly".

Figure 2 plots the ROC curves of all models for the 3 APR-DRGs with the highest in-hospital mortality rate. Supplementary Tables 2 and 3 display the model discrimination (c-statistic) and goodness-of-fit (R-square), respectively, along with 95% confidence intervals for all models and all APR-DRGs, ordered by their maximum value. Some examples of APR-DRGs with high discrimination are "843 – extensive 3rd degree or full thickness burns w/o skin graft" (Model 9 – 0.937, 95%CI 0.919-0.954; Model 11 – 0.953, 95%CI 0.935-0.969), "51 – viral meningitis" (Model 9 – 0.948, 95%CI 0.917-0.975; Model 11 – 0.933, 95%CI 0.896-0.968), "3 – bone marrow transplant" (Model 9 – 0.928, 95%CI 0.904-0.964; Model 11 – 0.929, 95%CI 0.902-0.956) or "650 – splenectomy" (Model 9 – 0.931, 95%CI 0.883-0.959; Model 11 – 0.934, 95%CI 0.898-0.958).

Discussion

In this study, we assessed the APR-DRGs' SOI and ROM as predictive models for in-hospital mortality, with both performing well when combined with age and sex. In fact, compared to CCI, ECI and some of their updates, both SOI and ROM (alongside age and sex) showed a statistically significantly better discrimination of in-hospital mortality in about half and a third of APR-DRGs, respectively. Additionally, between these two predictive models, SOI seemed to be the best performer for nearly 20% of APR-DRGs, while the opposite occurred in 2 out of 266 APR-DRGs. The model considering SOI, age and sex showed a good discrimination, with a c-statistic higher than 0.900 for 42.5% of APR-DRGs. Some particular APR-DRGs showed good discrimination, such as those related to burns, viral meningitis or specific transplants.

Romano and Chan have described model discrimination of SOI alone, ROM alone, each score with age and sex, as well as both scores together among patients with acute myocardial infarction in the US in 1991-1993 [23]. With models analogous to our models 9 (SOI, age and sex) and 11 (ROM, age and sex), they showed a c-statistic of 0.834 (95%CI 0.831-0.838) and 0.859 (95%CI 0.856-0.863), respectively. In the analogous APR-DRG "190 – acute myocardial infarction" in our sample, these values were of 0.814 (95%CI 0.806-0.823) and 0.824 (95%CI 0.819-0.832), respectively [23]. Although with statistically significant differences, these estimates are not too distant in both studies and both identify a tendency of the ROM model outperforming the SOI model for this particular set of patients. However, this tendency was not the most frequent in our study, as we found that the model with SOI, age and sex had a statistically significantly higher discrimination in 21.8% of APR-DRGs, compared to the model with ROM, age and sex, while the opposite only occurred in 2 out of 266 APR-DRGs. In fact, this is not in line with previous studies comparing SOI and ROM as predictive models for in-hospital mortality, showing that ROM

outperformed SOI, although considering wider samples and not only patients within each APR-DRG [22, 24]. Although not having access to the full algorithm to estimate SOI and ROM, the main differences are the default subclasses that are attributed to each ICD-9-CM secondary diagnosis code and how these default subclasses are gradually modified in each step according to the different combinations of patient variables [17, 26]. Besides this, the algorithm was developed using US hospital historical data. In fact, the differences found in the Portuguese setting are remarkable, with SOI outperforming ROM, as ROM was originally designed to predict the "risk of mortality". Furthermore, it is important to take into consideration that APR-DRGs' SOI and ROM, unlike the CCI and ECI, were developed to measure beyond the quantity of comorbidities, and while ROM subclass was specifically developed to predict mortality, the SOI subclass was developed to predict increased resource use due to the patient's acute problem, comorbidities and physiological decompensation [22].

Two studies have already showed that SOI and ROM models had better discrimination for predicting in-hospital mortality than the CCI or the ECI [22, 24]. In this study, we found similar results, showing that models with either SOI or ROM outperformed any of the CCI/ECI models in about half or a third of APR-DRGs, respectively. However, one important limitation of using APR-DRGs' SOI or ROM comparing to comorbidity measures is that the formers are APR-DRG specific and, thus, episodes of different APR-DRGs cannot be directly compared using either SOI or ROM, while they might be compared using CCI or ECI.

Furthermore, two studies have also combined SOI and ROM together and, despite showing better discrimination, estimates were not very different from ROM alone [23, 24].

Despite the lower performance of CCI or ECI, compared to the SOI or ROM models, such indices have proven to perform better in several contexts than more specific/physiology-based scores, although this is heterogeneous across diseases and contexts [27–30]. In addition to this, assessing and comparing SOI and ROM models with all these CCI/ECI models adds comprehensiveness to this study, being useful to those without access to APR-DRGs that use specific CCI/ECI models.

These type of scores or indices can be used for several purposes besides predicting mortality such as predicting hospital readmissions [31]. Nevertheless, using them only for predicting (in-hospital) mortality can be (and has been) useful for several purposes, including benchmarking hospitals or organization models, analyzing epidemiological patterns or trends, planning and allocating resources, informing patient prognosis, improving/tailoring treatments or even increasing clinical trials comparability [1, 32–36].

These scores or indices rely on the clinical coding quality on administrative databases, which can be compared and show closer performances to chart review comorbidity measures [37–39]. DRG payment schemes also rely on administrative databases, in which the gathered information is usually coded after patients' discharge. In fact, if and when we achieve high-quality real-time or automated clinical coding, data will be more reliable and will have more clinical applications, including in-hospital mortality prediction. Actually, there is increasing research on using artificial intelligence for clinical coding automation, namely with machine learning and natural language processing techniques [40–42]. Moreover, additional clinical information not usually coded into such administrative databases is being more and more combined with artificial intelligence techniques for predicting mortality [43–46].

One of the limitations of this study may be the exclusion criteria, which comprised episodes transferred to another hospital, patients "leaving against medical advice" and episodes with a length of stay equal to 0, which we believed that were the most unbiased options considering the study aims. Regarding the latter, we also performed robustness checks, including those episodes, and differences were small. Another important limitation is the underlying clinical coding data, which can have several pitfalls as all DRG-related administrative databases [47]. Moreover, we opted for selecting only ICD-9-CM data as the transition from ICD-9-CM to ICD-10-CM in Portugal started in October 2016 for pilot hospitals and in the beginning of 2017 for the rest and this might bias the results [48].

In conclusion, besides the comorbidity indices that are currently used to predict in-hospital mortality prediction, APR-DRGs' SOI and ROM can also be considered for this purpose. However, APR-DRGs' SOI and ROM are DRG-specific, and these subclasses are not directly comparable between DRGs. Thus, its use for risk-adjustment in broader settings (e.g., overall in-hospital mortality) must be considered with caution. We found that SOI or ROM, combined with age and sex, perform better than traditional and more widely used comorbidity indices such as CCI or ECI. This is particular important as these scores are derived from the same administrative databases, although APR-DRGs' SOI and ROM data can be readily available in several or most institutions where APR-DRG is adopted. Our study also concluded that, although ROM is specifically produced to predict mortality, SOI performed better than ROM in almost a quarter of APR-DRGs. In the future, these findings can be clinically important with the real-time or automated clinical coding. Meanwhile, the findings can be quite helpful for hospital or organizational models benchmarking or epidemiological analysis.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s10916-022-01805-3.

Funding This study received no funding.

Declarations

Conflicts of interest The authors declare no conflicts of interest.

References

1. Minne L, Ludikhuize J, de Rooij SE, Abu-Hanna A. Characterizing predictive models of mortality for older adults and their validation for use in clinical practice. J Am Geriatr Soc. 2011 Jun;59(6):1110-5.
2. Larsson J, Itenov TS, Bestle MH. Risk prediction models for mortality in patients with ventilator-associated pneumonia: A systematic review and meta-analysis. J Crit Care. 2017 Feb;37:112-118.
3. Sharma, N., Schwendimann, R., Endrich, O., Ausserhofer, D., & Simon, M. (2021). Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data. BMC health services research, 21(1), 13.
4. 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: 373–383.
5. Elixhauser A, Steiner C, Robert Harris D, Coffey RM. Comorbidity measures for use with administrative data. Med Care 1998; 36: 8–27.
6. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005; 43: 1130–1139.
7. Quan H, Bing L, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. Am J Epidemiol 2011; 173: 676–682.
8. 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. Med Care 2009; 47: 626–633.
9. Moore BJ, White S, Washington R, Coenen N, Elixhauser A. Identifying increased risk of readmission and in-hospital mortality using hospital administrative data. Med Care 2017; 55: 698–705.
10. Schneeweiss S, Wang PS, Avorn J, Glynn RJ. Improved comorbidity adjustment for predicting mortality in Medicare populations. Health Serv Res. 2003; 38(4):1103–1120.
11. Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, Griffith SD. A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. Med Care. 2015; 53(4):374–379.
12. Fetter RB, Thompson JD, Mills RE. A system for cost and reimbursement control in hospitals. Yale J Biol Med 1976; 49: 123–136.
13. Busse R, Geissler A, Quentin W, et al. Diagnosis Related Groups in Europe. Moving towards Transparency, Efficiency and Quality in Hospitals. Copenhagen: Open University Press. 2011.
14. Schreyogg J, Stargardt T, Tiemann O, Busse R. Methods to determine reimbursement rates for diagnosis related groups (DRG): a comparison of nine European countries. Health Care Management Sci 2006; 9(3): 215–223.
15. Mathauer I, Wittenbecher F. Hospital payment systems based on diagnosis-related groups: Experiences in low- and middle-income countries. Bull World Health Organ 2013; 91: 746–756.
16. Mihailovic, N., Kocic, S., & Jakovljevic, M. Review of Diagnosis-Related Group-Based Financing of Hospital Care. Health services research and managerial epidemiology, 3, 2333392816647892. 2016.
17. Averill R, Goldfield N, et al. All Patient Refined Diagnosis Related Groups (APR-DRGs) Version 31.0: Methodology Overview [Internet]. Available from: https://hcup-us.ahrq.gov/db/nation/nis/grp031_aprdrg_meth_ovrview.pdf [Accessed in Jun 15 2021].
18. Dewilde, S, Annemans, L, Pincé, H, Thijs, V. Hospital financing of ischemic stroke: determinants of funding and usefulness of DRG subcategories based on severity of illness. BMC Health Serv Res. 2018; 18(1): 356.
19. Nante N, De Marco MF, Balzi D, Addari P, Buiatti E. Prediction of mortality for congestive heart failure patients: results from different wards of an Italian teaching hospital. Eur J Epidemiol. 2000; 16: 1017–21.
20. Iezzoni LI, Shwartz M, Ash AS, Mackiernan YD. Predicting in-hospital mortality for stroke patients: results differ across severity-measurement methods. Med Decis Making. 1996; 16(4): 348–356.
21. Baram D, Daroowalla F, Garcia R, Zhang G, Chen JJ, Healy E, Riaz SA, Richman P. Use of the All Patient Refined-Diagnosis Related Group (APR-DRG) Risk of Mortality Score as a Severity Adjustor in the Medical ICU. Clin Med Circ Respirat Pulm Med. 2008; 2: 19–25.
22. McCormick PJ, Lin HM, Deiner SG, Levin MA. Validation of the All Patient Refined Diagnosis Related Group (APR-DRG) Risk of Mortality and Severity of Illness Modifiers as a Measure of Perioperative Risk. J Med Syst. 2018; 42(5): 81.
23. Romano PS, Chan BK. Risk-adjusting acute myocardial infarction mortality: are APR-DRGs the right tool? Health Serv Res. 2000 Mar;34(7):1469–89.
24. Liu J, Larson E, Hessels A, Cohen B, Zachariah P, Caplan D, Shang J. Comparison of Measures to Predict Mortality and Length of Stay in Hospitalized Patients. Nurs Res. 2019; 68(3): 200–209.
25. DiCiccio TJ and Efron B. Bootstrap Confidence Intervals. Statistical Science 1996; 11: 189.
26. Souza J, Santos JV, Canedo VB, Betanzos A, Alves D, Freitas A. Importance of coding co-morbidities for APR-DRG assignment: Focus on cardiovascular and respiratory diseases. Health Inf Manag 2020; 49: 47–57.
27. Bahlis LF, Diogo LP, Fuchs SC. Charlson Comorbidity Index and other predictors of in-hospital mortality among adults with community-acquired pneumonia. J Bras Pneumol. 2021; 47(1): e20200257.
28. 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.
29. 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.
30. Bennett CE, Wright RS, Jentzer J, Gajic O, Murphree DH, Murphy JG, Mankad SV, Wiley BM, Bell MR, Barsness GW. Severity of illness assessment with application of the APACHE IV predicted mortality and outcome trends analysis in an academic cardiac intensive care unit. J Crit Care. 2019; 50: 242–246.
31. Wang H, Johnson C, Robinson RD, Nejtek VA, Schrader CD, Leuck J, Umejiego J, Trop A, Delaney KA, Zenarosa NR. Roles of disease severity and post-discharge outpatient visits as predictors of hospital readmissions. BMC Health Serv Res. 2016; 16(1): 564.
32. Jokinen JJ. Why do we have to predict mortality rates? Acta Anaesthesiol Scand. 2011; 55(3): 255–256.

33. Mascha EJ, Mazo V. Predicting In-hospital Postoperative Mortality for the Practitioner: Beyond the Numbers. Anesthesiology. 2016; 124(3): 523–525.
34. El Amrani M, Clement G, Lenne X, Rogosnitzky M, Theis D, Pruvot FR, Zerbib P. The Impact of Hospital Volume and Charlson Score on Postoperative Mortality of Proctectomy for Rectal Cancer: A Nationwide Study of 45,569 Patients. Ann Surg. 2018; 268(5): 854–860.
35. Jarman B, Pieter D, van der Veen AA, Kool RB, Aylin P, Bottle A, Westert GP, Jones S. The hospital standardised mortality ratio: a powerful tool for Dutch hospitals to assess their quality of care? Qual Saf Health Care. 2010; 19(1): 9–13.
36. Bhandari S, Abdul MKM, Hollabaugh W, Sharma K, Evans DB, Guda N. Decreased trend in hospital mortality from pancreatic cancer despite increase in number of hospital admissions. PLoS One. 2018; 13(7): e0199909.
37. Stavem K, Hoel H, Skjaker SA, Haagensen R. Charlson comorbidity index derived from chart review or administrative data: agreement and prediction of mortality in intensive care patients. Clin Epidemiol. 2017; 9: 311–320.
38. Weir RE Jr, Lyttle CS, Meltzer DO, Dong TS, Ruhnke GW. The Relative Ability of Comorbidity Ascertainment Methodologies to Predict In-Hospital Mortality Among Hospitalized Community-acquired Pneumonia Patients. Med Care. 2018; 56(11): 950–955.
39. Peng M, Southern DA, Williamson T, Quan H. Under-coding of secondary conditions in coded hospital health data: Impact of co-existing conditions, death status and number of codes in a record. Health Informatics J. 2017; 23(4): 260–267.
40. Ramalho A., Souza J., Freitas A. The Use of Artificial Intelligence for Clinical Coding Automation: A Bibliometric Analysis. In: Dong Y., Herrera-Viedma E., Matsui K., Omatsu S., González Briones A., Rodríguez González S. (eds) Distributed Computing and Artificial Intelligence, 17th International Conference. DCAI 2020. Advances in Intelligent Systems and Computing, vol 1237. Springer, Cham. 2021.
41. Stanfill MH, Williams M, Fenton SH, Jenders RA, Hersh WR. A systematic literature review of automated clinical coding and classification systems. J Am Med Inform Assoc. 2010; 17(6): 646–651.
42. Catling F, Spithourakis GP, Riedel S. Towards automated clinical coding. Int J Med Inform. 2018; 120: 50–61.
43. Mpanya D, Celik T, Klug E, Ntsinjana H. Machine learning and statistical methods for predicting mortality in heart failure. Heart Fail Rev. 2021 May;26(3):545-552. https://doi.org/10.1007/s10741-020-10052-y. Epub 2020 Nov 9. PMID: 33169338.
44. Weissman GE, Hubbard RA, Ungar LH, Harhay MO, Greene CS, Himes BE, Halpern SD. Inclusion of Unstructured Clinical Text Improves Early Prediction of Death or Prolonged ICU Stay. Crit Care Med. 2018; 46(7): 1125–1132.
45. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, Liu PJ, Liu X, Marcus J, Sun M, Sundberg P, Yee H, Zhang K, Zhang Y, Flores G, Duggan GE, Irvine J, Le Q, Litsch K, Mossin A, Tansuwan J, Wang D, Wexler J, Wilson J, Ludwig D, Volchenboum SL, Chou K, Pearson M, Madabushi S, Shah NH, Butte AJ, Howell MD, Cui C, Corrado GS, Dean J. Scalable and accurate deep learning with electronic health records. NPJ Digit Med. 2018; 1: 18.
46. Syed M, Syed S, Sexton K, Greer ML, Zozus M, Bhattacharyya S, Syed F, Prior F. Deep Learning Methods to Predict Mortality in COVID-19 Patients: A Rapid Scoping Review. Stud Health Technol Inform. 2021; 281: 799–803.
47. Alonso V, Santos JV, Pinto M, Ferreira J, Lema I, Lopes F, Freitas A. Health records as the basis of clinical coding: Is the quality adequate? A qualitative study of medical coders' perceptions. Health Inf Manag. 2020; 49(1): 28–37.
48. Santos JV, Novo R, Souza J, Lopes F, Freitas A. Transition from ICD-9-CM to ICD-10-CM/PCS in Portugal: An heterogeneous implementation with potential data implications. Health Inf Manag. 2021 [epub ahead of print].