Journal of Clinical Epidemiology 116 (2019) 62−71

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

A new population-based risk stratification tool was developed and validated for predicting mortality, hospital admissions, and health care costs

Federico Rea[a,b,*], Giovanni Corrao[a,b], Monica Ludergnani[c], Luigi Cajazzo[d], Luca Merlino[e]

[a]National Centre for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy
[b]Laboratory of Healthcare Research and Pharmacoepidemiology, Unit of Biostatistics, Epidemiology and Public Health, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy
[c]Lombardia Informatica S.p.a, Milan, Italy
[d]Regional Welfare Service, Lombardy Region, Milan, Italy
[e]Epidemiologic Observatory, Regional Welfare Service, Lombardy Region, Milan, Italy

Accepted 23 August 2019; Published online 28 August 2019

Abstract

Objectives: The aim of this study was to develop a new population-based risk stratification tool (Chronic Related Score [CReSc]) for predicting 5-year mortality and other outcomes.

Study Design and Setting: The score included 31 conditions selected from a list of 65 candidates whose weights were assigned according to the Cox model coefficients. The model was built from a sample of 5.4 million National Health Service (NHS) beneficiaries from the Italian Lombardy Region and applied to the remaining 2.7 million NHS beneficiaries. Predictive performance was assessed by discrimination and calibration. CReSc ability in predicting secondary endpoints (i.e., hospital admissions and health care costs) was investigated. Finally, the relationship between CReSc and income was considered.

Results: Among individuals aged 50−85 years, CReSc performance showed (1) an area under the receiver operating characteristic curve of 0.730, (2) an improved reclassification from 44% to 52% with respect to other scores, and (3) a remarkable calibration. A trend toward increasing rates of all the considered endpoints as CReSc increases was observed. Compared with individuals on low−intermediate income, NHS beneficiaries on high income showed better CReSc profile.

Conclusion: We developed a risk stratification tool able to predict mortality, costs, and hospital admissions. The application of CReSc may generate clinically and operationally important effects.

Keywords: Comorbidity; Health care utilization database; Prognostic score; Population-based study; Record linkage; Risk stratification

1. Introduction

The aging population with rising prevalence of chronic conditions makes unprecedented demands on health care services [1,2]. Patients with chronic conditions are more likely to experience hospital admissions for potentially avoidable causes, resulting in a "triple fail event" of suboptimal health outcomes, significant health costs, and poor patient experience [3,4].

Through detecting high-risk individuals who require more careful care [5], risk stratification models are expected to improve patient outcomes and provide economic value to health care [6,7]. This proactive targeting of services for people at defined risk has retained prominence by the Italian Health Ministry since, notably with the recently adopted national plan of chronicity, efforts to

Conflicts of interest: G.C. received research support from the European Community (EC), the Italian Agency of Drug (AIFA), and the Italian Ministry of Education, University and Research (MIUR). He took part in a variety of projects that were funded by pharmaceutical companies (i.e., Novartis, GSK, Roche, AMGEN, and BMS). He also received honoraria as member of Advisory Board from Roche. For the remaining authors, nothing was declared.

Funding: This study was funded by grants from the Italian Ministry of the Education, University and Research, Italy ("PRIN" 2017, project 2017728JPK). Italian Ministry of the Education, University and Research had no role in the design of the study, the collection, analysis, and interpretation of the data, as well as the writing of the manuscript.

* Corresponding author. Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, Edificio U7, 20126 Milano, Italy. Tel.: +39-02-64485859; fax: +39-02-64485899.

E-mail address: f.rea@campus.unimib.it (F. Rea).

https://doi.org/10.1016/j.jclinepi.2019.08.009

What is new?

Key findings

- We developed and validated a population-based risk stratification tool (Chronic-Related Score) derived from several sources of health care data, useful for predicting 5-year mortality and other outcomes.
- The Chronic-Related Score significantly outperformed the Charlson Comorbidity Score and Multisource Comorbidity Score in terms of discriminatory power and net reclassification improvement. In addition, Chronic-Related Score showed a remarkable calibration.

What this adds to what was known?

- Our score is the first tool able to combining multiple sources and predicting health outcomes in the general population.

What is the implication and what should change now?

- The Chronic-Related Score can represent a useful tool for risk adjustment in clinical and epidemiological studies.
- The introduction of our risk stratification tool may generate clinically and operationally important effects.

introduce integrated care mainly addressed to high-risk individuals have been recommended (http://www.salute.gov.it/imgs/C_17_pubblicazioni_2584_allegato.pdf).

In estimating individual risk scores, models typically include predictors relating to the past use of health care, diagnoses, and medications [8–14]. However, most of the available scores have been developed from hospital-based surveys reviewing inpatients medical records, and only later they were adapted for risk stratification of whoever beneficiary of the health system [14–17]. A simple population-based comorbidity score (i.e., Multisource Comorbidity Score [MCS]) has been recently validated by our group in the setting of the Italian National Health System [18]. Because MCS was developed to be applied everywhere, only data on inpatient diagnoses and outpatient drug prescriptions were used. However, other health care utilization data, although not ubiquitously collected, may contribute to explain relevant clinical and economic outcomes.

Recognizing such challenges and opportunities, welfare authority from the Lombardy Region of Italy developed an ''evidence-based'' framework involving the building of algorithms able of capturing patients who suffer from 65 chronic diseases and conditions through all the available health care utilization databases, thus realizing the so-called Chronic Related Groups (CReG).

Herein, taking advantage from several sources of health care data routinely collected for National Health Service (NHS) beneficiaries from Lombardy, we developed and validated a score based on the CReG list, which we just called Chronic-Related Score (CReSc). We investigated the association of the CReSc with 5-year mortality, hospital admissions, and health care costs. In addition, the relationship between CReSc distribution and household income was investigated.

2. Methods

2.1. Healthcare utilization data from Lombardy

The data used for the present study were retrieved from the health care utilization databases of Lombardy, a region of Italy that accounts for about 16% of its population. All Italian citizens have equal access to health care services as part of the NHS. The NHS guarantees partly or entirely free of charge access to a number of health care services (the so-called essential levels of assistance) to all Italian citizens. In Lombardy, this has been associated since 1997 with an automated system of health care utilization databases to collect a variety of information, including the (1) archive of NHS beneficiaries, reporting demographic and administrative data, other than the dates in which the condition of NHS beneficiary started (because he/she was born or immigrated) or stopped (because he/she died or emigrated); (2) hospital discharge database reporting information about primary diagnosis, coexisting conditions, and provided procedures (coded according to [a] the International Classification of Diseases, Ninth Revision, Clinical Modification, classification system, http://icd9.chrisendres.com/; and [b] the official Italian diagnosis-related group, http://www.gazzettaufficiale.it/eli/gu/2010/01/05/3/sg/pdf); (3) drug prescription database providing information on all the drugs reimbursed by the NHS (coded according to the Anatomical Therapeutic Chemical classification system, https://www.whocc.no/atc_ddd_index/); (4) outpatient database, including visits in specialist ambulatories and diagnostic laboratories accredited from the NHS (coded according to the regional outpatients services coding, http://normativasan.servizirl.it/port/GetNormativaFile?fileName=13400_prest_AMB_da_01%20LUGLIO%202018%20Pubblicato.xls); and (5) copayment exception database, including exception for chronic disease (coded according to the national exceptions coding, http://www.trovanorme.salute.gov.it/norme/renderPdf.spring?seriegu=SG&datagu=18/03/2017&redaz=17A02015&artp=13&art=1&subart=1&subart1=10&vers=1&prog=001). Costs associated with the provided health care services were measured from the NHS perspective using the amount that the Regional Health Authority reimbursed to health providers.

As a unique identification code was used for all databases, their record linkage allowed searching out the complete care pathway of NHS beneficiaries. To preserve privacy, identification codes were automatically converted into pseudoanonymized codes, and the inverse process was prevented by the deletion of the conversion table.

2.2. Score development

A list of 65 chronic conditions (i.e., the CReG list, which is reported in the Supplementary Table S1) was carefully chosen by a regional working group appointed for identifying those conditions, which mostly affect expenditure of regional health authority. For each of the 65 conditions, an algorithm was developed for capturing patients who suffer from it through the above-mentioned databases. With the aim of selecting conditions independently able to predict 5-year mortality (i.e., the main outcome of interest), we proceeded as follows.

First, two out of three of the 8.1 million citizens aged 18 years or older who in the year 2013 were beneficiaries of the Lombardy health system from at least 2 years (i.e., almost 5.4 million citizens) was randomly selected to form the so-called training (derivation) set. These patients were followed until the earliest date between death and censoring (emigration or December 31, 2017). Second, the Cox proportional hazard regression model was fitted to compute the hazard ratios estimating the relationship between the selected covariates and the time of death. Covariates included in the model were gender, age (at January 1, 2013), and the 65 candidate predictors. The latter entered into the model as dichotomous variables, with value 1 or 0 according to whether the specific condition was or was not recorded at least once within 2 years before baseline (2011−2012). Third, the least absolute shrinkage and selection operator (LASSO) method was applied for selecting the diseases/conditions able to independently predicting 5-year mortality [19]. LASSO selects variables correlated to the measured outcome by shrinking coefficients weights, down to zero for the ones not correlated to outcome. Finally, the coefficients estimated from the model were used for assigning a weight at each selected covariate. In particular, a weight was assigned to each coefficient by multiplying it by 10 and rounding it to the nearest whole number [20]. The weights thus obtained were then summed to produce a total aggregate score. To simplify the system, that is, with the aim of accounting for excessive heterogeneity of the total aggregate score, the latter was categorized by assigning increasing values of 0, 1, 2, 3, and 4 to the categories of the aggregate score of 0, 1−10, 11−20, 21−30, and ≥31, respectively. The so obtained index was denoted CReSc.

2.3. Model performance

Performance of CReSc was explored with respect to other prognostic scores such as the Charlson Comorbidity Index (CCI) and the Multisource Comorbidity Score (MCS). With this aim, the corresponding weights were applied to the so-called validation set consisting of the beneficiaries of NHS who did not enter into the training set (i.e., 2.7 million).

Predictive performance was assessed through discrimination and calibration. Discrimination indicates how well the model can distinguish individuals with the outcome from those without the outcome. Two approaches were used for assessing discrimination: (1) discriminatory powers were compared by the receiver operating characteristic (ROC) curves and the corresponding area under the ROC curves (AUCs) [21]; (2) the net reclassification improvement (NRI) was calculated for assessing the ''net'' number of individuals correctly reclassified using CReSc over a comparator index (i.e., CCI or MCS) [22].

Calibration ascertains the concordance between the model's predictions and observed outcomes, which we evaluated using a calibration plot. The plot displays predicted versus observed 5-year survival probabilities for increasing the predicted risk. Ideally, the plot follows a 45° line, showing that the predicted risks are equal to the observed outcome frequencies. We assessed the difference in predicted and observed frequency in the total cohort, indicating the extent to which predictions are systematically too high or too low (referred to as calibration-in-the-large), and the recalibration slope, reflecting the slope of the calibration plot and ideally equal to 1 [23]. Finally, the Hosmer−Lemeshow goodness-of-fit test modified by Yu et al. [24] was used for testing the null hypothesis of agreement between observed and predicted survival probabilities.

Other than for the entire sample, performance of CReSc was calculated according to age (three classes of 18 until 50 years, 51 until 85 years, and 86 years or older) and gender.

2.4. Secondary outcomes

Secondary analyses were performed for verifying the CReSc robustness in predicting outcomes other than the cumulative 5-year all-cause mortality. With this aim, cumulative health care costs, rates of hospital admissions, and cumulative days of hospital stays were calculated along the CReSc increasing categories for the entire 5-year time window. Cumulative health care costs were calculated by means of the Bang & Tsiatis estimator [25], a method that takes into account censored cost data. Finally, the rates of hospital admissions and cumulative days of hospital stays, both expressed as average number every 1,000 person-years (PY), were considered.

2.5. Exploring the effect of income on the CReSc

Because of the well-known association between income inequality and health [26] and with the aim to verify

whether income affects CReSc profiles, data were further processed as follows.

First, we calculated the individual income from the household taxable income data (regional tax registry), according to the above-reported rules of pseudoanonymization. To take into account the household size and composition, we used the formula that is frequently applied in income inequalities studies, which consists of dividing the household income by the square root of the number of household members [27]. When counting the number of household members, the first adult of a family was weighted as 1.0, other adults as 0.7, and children less than 18 years old as 0.5. The equivalent income measures the net income per year that is available for one person.

Second, we classified NHS beneficiaries according to their income. Low- and high-income categories were defined according to 10th and 90th percentiles of equivalent income distribution, being intermediate-income category otherwise. Because the income is expected to be related to age and gender, we used the equivalent income distribution for each stratum of gender and 5-year intervals of age for categorizing individuals in low, intermediate, and high income.

Third, with the aim of assessing whether income categories affected the CReSc value, the ridit analysis was used. Ridit transformation is a method based on the assumption that the ordered categorical variable approximates an underlying continuous variable [28,29]. In the present study, within each stratum of gender and 5-year intervals of age, we ridit transformed the CReSc so to calculate the mean ridit for NHS beneficiaries on low and high income with respect to those on intermediate income. Mean ridit measures the probability that a randomly selected NHS beneficiary among those on low or high income had higher CReSc value than a randomly selected NHS beneficiary among those on intermediate income [30−32]. The resulting ridit score represents a probability that ranges from a minimum of 0 to a maximum of 1, with a value of 0.5 assigned to the reference category. For example, whether a trend toward worsening CReSc as income get worse, then ridit values higher and lower than 0.5 would be respectively expected among NHS beneficiaries on low and high income. The statistical significance of differences in mean ridit was evaluated with a Z-test.

Finally, since the CReSc score could have different performances according to the income categories (e.g., because NHS beneficiaries with low income could have worse access to services), we took some caution by comparing the ROC curves, and the corresponding AUCs, among categories of equivalent income. Of course, overlapping ROC curves and similar AUC values are expected. Otherwise, we should interpret the ridit values among the income categories as an effect of the different performance between the income categories, rather than the income action in differentiating the level of chronicity.

3. Results

3.1. Chronic-Related Score

The 31 conditions significantly contributing to the CReSc, the corresponding codes, jointly with a schematic explanation on how the score must be calculated, are reported in the supplementary material (Supplementary Table S2). Almost 30% of NHS beneficiaries had at least a condition contributing to the CReSc. Table 1 shows that Alzheimer's disease, dementia, and heart failure most contributed to the total aggregate score. As expected, with respect to the other listed conditions, hypertension, type II diabetes, and no arrhythmia-induced cardiomyopathy had higher prevalence rates, whereas dialytic treatment, HIV/AIDS, and acromegaly and gigantism had higher per-capita health care cost. By considering prognosis, prevalence, and costs, in a unique proxy of disease burden, active neoplasia, heart failure, and liver cirrhosis showed a higher impact in our setting.

Fig. 1 shows that the worst clinical status expected for men and elderly with respect to women and young people was clearly caught through the CReSc, being the prevalence of NHS beneficiaries with at least a chronic disease contributing to the score progressively increased with age both in women (from 6.6% to 74.6%) and even more in men (from 7.7% to 79.1%).

3.2. CReSc performance

AUC of CReSc, MCS, and CCI had values of 0.793, 0.761, and 0.623, respectively (Supplementary Fig. S1). It should be emphasized that because the very large sample size, AUC and confidence limits practically coincided so that confidence intervals were not reported. Fig. 2 shows that our score had good performance among both women and men belonging to the intermediate age category (50 until 85 years) being the corresponding AUC 0.730 and 0.731. Conversely, among younger and even more among older beneficiaries, our score had a poor predictive ability (being the corresponding AUC 0.687 and 0.570 in women and 0.637 and 0.627 in men). For this reason, further analyses were performed by excluding NHS beneficiaries younger than 50 years and those older than 85 years.

Performance analyses using NRI showed that CReSc significantly improved the net 5-year mortality reclassification. In particular, CReSc improved MCS classification of 44% (95% CI 43−45%) both in men and women. Of interest, the improvement mainly concerned sensitivity (25%) than specificity (18%). As expected, CReSc even more improved the classification of CCI, being the corresponding NRI 50% (49−51%) and 52% (51−53%) in men and women respectively.

Fig. 3 shows that there was a good agreement between the observed and the predicted survival probabilities, with values of the calibration-in-the-large close to the ideal value of 0 (−0.03) and values of the recalibration slopes close to the ideal value of 1 (1.03). Adequate goodness-

Table 1. Weights, prevalence, and burden of conditions contributing to the Chronic-Related Score (CReSc)

#	Disease/condition[a]	CReSc weight[b]	Prevalence rate (%)	Per-capita annual cost (€)	Burden ranking[c]
1	Alzheimer disease	22	0.26	9,026	29
2	Dementia	20	0.17	9,684	23
3	Heart failure	18	1.80	20,345	2
4	Neoplasia, active	14	2.10	20,828	1
5	Parkinson diseases	14	0.40	18,246	6
6	Dialysis	14	0.09	140,824	10
7	Arrhythmic myocardiopathy	11	2.15	14,650	7
8	Liver cirrhosis	10	0.30	27,862	3
9	Respiratory insufficiency/oxygen therapy	10	0.10	21,971	11
10	Hypertension	9	15.47	9,203	18
11	Ischemic cardiopathy	9	2.47	15,078	4
12	Chronic obstructive pulmonary	9	1.86	17,187	5
13	Cerebral vasculopathy	9	1.04	15,430	9
14	Immune hemolytic anaemias	9	0.01	23,377	28
15	Neoplasia, follow-up	8	1.21	14,746	12
16	Diseases of the nervous system and the sense organs	7	0.10	14,609	21
17	Cardiomyopathy (no arrhythmia induced)	6	2.67	13,977	13
18	Chronic kidney failure	6	0.75	24,513	8
19	Acromegaly and gigantism	6	0.02	52,394	24
20	Type II diabetes mellitus	5	5.58	13,659	15
21	Epilepsy	5	0.44	12,395	25
22	Venous vasculopathy	5	0.22	14,927	19
23	Systemic sclerosis	5	0.06	28,637	17
24	Chronic pancreatitis	5	0.05	25,327	22
25	Myasthenia gravis	5	0.03	18,299	27
26	Arterial vasculopathy	4	0.56	22,986	14
27	Type II diabetes mellitus, complicated	4	0.40	28,095	16
28	Addison's diseases	4	0.02	12,909	31
29	HIV positive and full-blown AIDS	3	0.30	54,036	20
30	Chronic hepatitis	1	0.94	18,217	26
31	Valvular cardiopathy	1	0.45	13,862	30
	Beneficiaries with at least one selected disease/condition	-	29.94	-	

[a] Disease/condition selected as independent predictor of 5-year mortality.
[b] Weight calculated by multiplying by 10 the specific coefficient of the survival model and rounding it to the nearest whole number.
[c] Ranking obtained by multiplying ranking of CReSc weight, ranking of prevalence rate, and ranking of per-capita annual cost of each selected disease/condition.

of-fit was also confirmed by the modified Hosmer–Lemeshow test, according to which the null hypothesis of agreement between observed and predicted frequencies could not be rejected either in women or in men.

3.3. CReSc predictability

A clear positive trend toward increasing rates of all the considered outcomes as CReSc increases was observed (Fig. 4). In particular, with respect to NHS beneficiaries with the lowest score (CReSc = 0), those on the highest score (CReSc = 4) had 1-year risk of death, 5-year risk of death, 5-year health care costs, rate of hospital admissions, and rate of cumulative hospital stay, respectively, 36-fold (from 0.43% to 15.5%), 15-fold (from 3.43% to 52.6.21%), 6-fold (from €4,236 to €26,882), 9-fold (from 113 to 761 hospital admissions every 1,000 PY), and 7-fold (from 849 to 8,839 days of stay every 1,000 PY) higher.

3.4. CReSc and income relationship

AUC for NHS beneficiaries with low, intermediate, and high income had values of 0.726, 0.723, and 0.727, respectively (Fig. 5, left box). Starting from the age category of

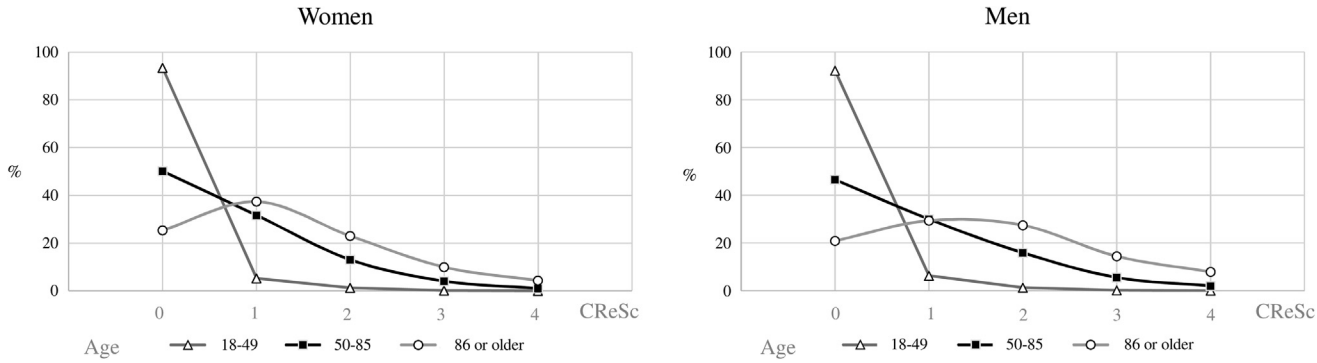

Fig. 1. Chronic-Related Score (CReSc) distribution among National Health Services beneficiaries of Lombardy, Italy, according to their gender and age category.

61–65 years, a trend toward ameliorating CReSc profile with income category augmenting was observed both in women and in men (Fig. 5, right boxes). Compared with individuals on low-intermediate income, NHS beneficiaries on the highest income always showed better CReSc profile irrespectively from age and gender.

4. Discussion

Our study shows that a score based on health care utilization data currently used for managing NHS from Italian Lombardy health authority is able to stratify NHS beneficiaries according to their short-term (1-year) and long-term (5-year) outcomes such as mortality, health care costs, and hospital admissions. Our score significantly improved discriminatory power and net reclassification of the CCI, undoubtedly the most worldwide used comorbidity score [15]. In addition, because our score used much more information than the MCS, CReSc outperformed the previous population-based comorbidity score validated by our group in which only data on inpatient diagnoses and outpatient drug prescriptions were used [18]. Compared with MCS, and even more CCI, CReSc showed better ability to identify NHS beneficiaries at higher risk (so allowing better identification of patients who need medical care), and even more to exclude those at low risk of experiencing adverse outcomes. It follows that CReSc may be useful to epidemiologists, clinicians, and policy-makers who now have a tool for risk adjustment, management, and stratification characterized by improved performances with respect to the available scores.

Our study provides the following additional results. First, among the 31 conditions considered, Alzheimer's

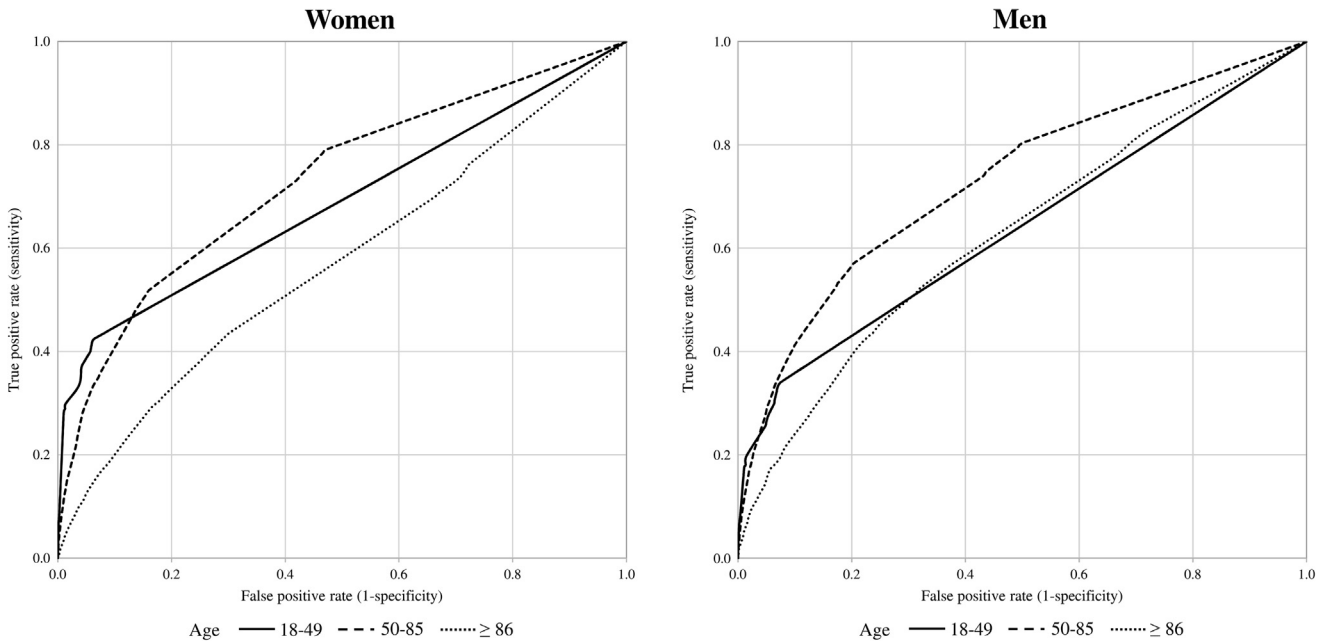

Fig. 2. Receiver Operating Characteristics curves comparing discriminant power of Chronic-Related Score according to age strata and gender in predicting 5-year mortality among National Health Services beneficiaries of Lombardy, Italy.

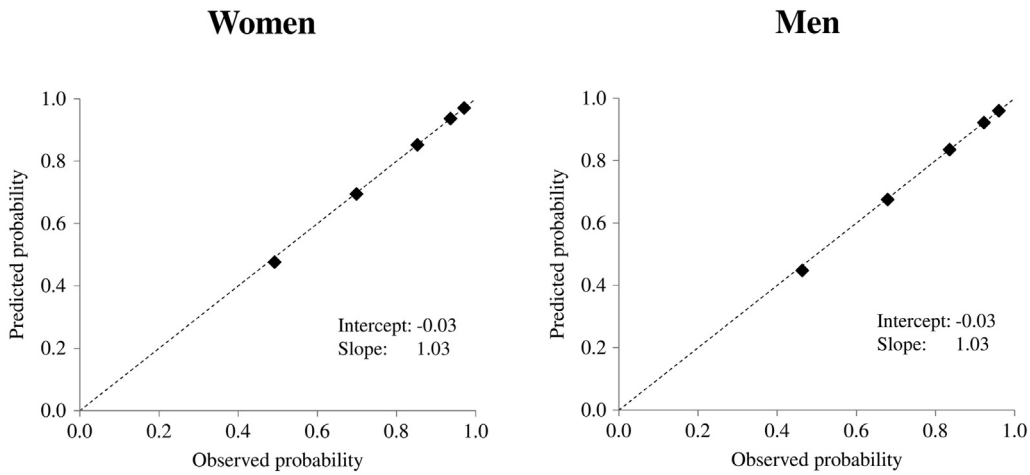

Fig. 3. Calibration plots comparing observed and predicted 5-year survival probabilities of the Chronic-Related Score according to gender of National Health Services beneficiaries of Lombardy, Italy.

disease, dementia, and heart failure most contributed to the total aggregate score. Of interest, our data confirmed that in our setting, most persons live with hypertension and/or type II diabetes (roughly one in five NHS beneficiaries live with one or both the conditions), and that dialytic treatment and positivity to HIV and full-blown AIDS had high financial burden (being health care costs associated with these conditions €140,000 and €54,000, respectively). By including prognostic weight, prevalence, and costs in a unique proxy of disease burden, active neoplasia, heart failure, and liver cirrhosis showed a higher impact in our population.

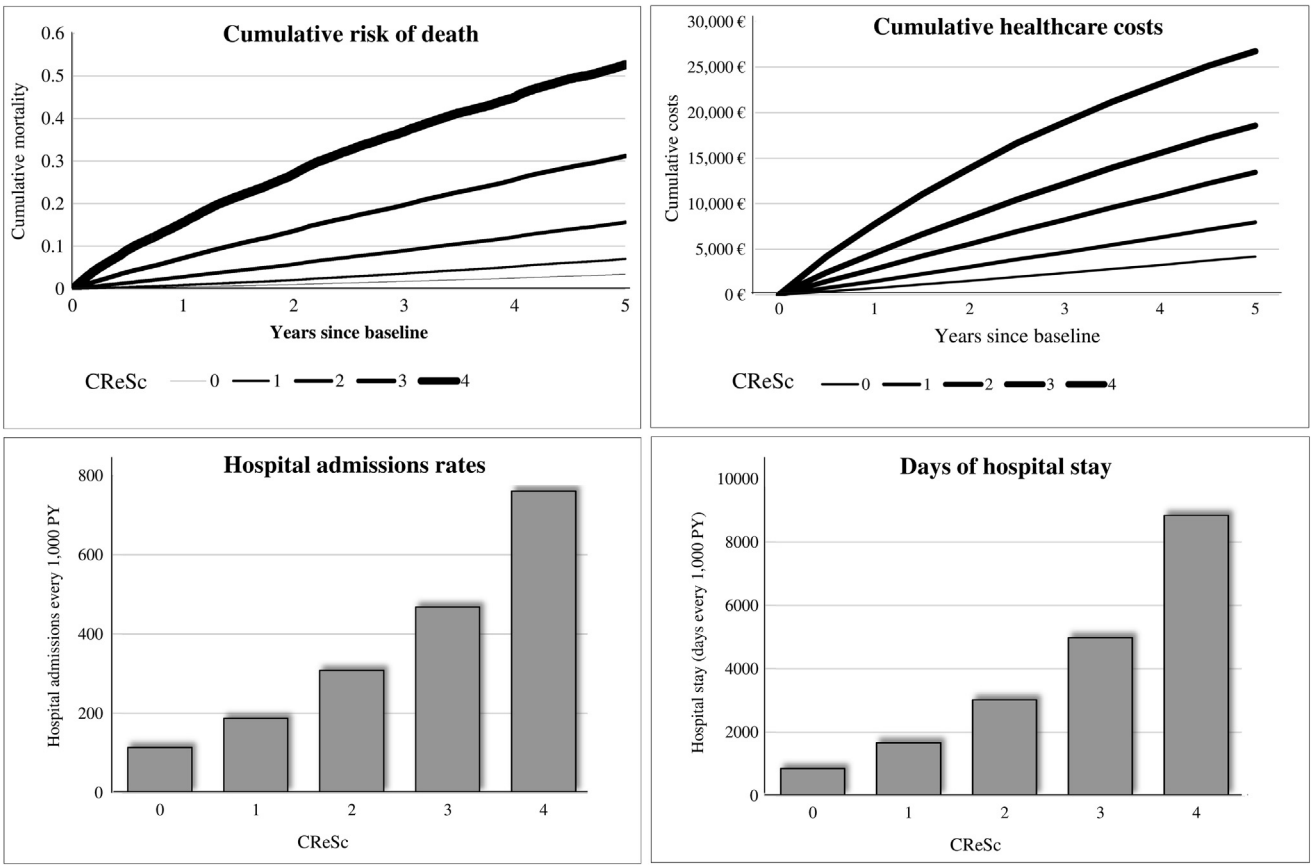

Fig. 4. Five-year cumulative mortality and health care costs and rates of hospital admission and days of hospital stay, according to the Chronic-Related Score distribution among NHS beneficiaries of Lombardy, Italy.

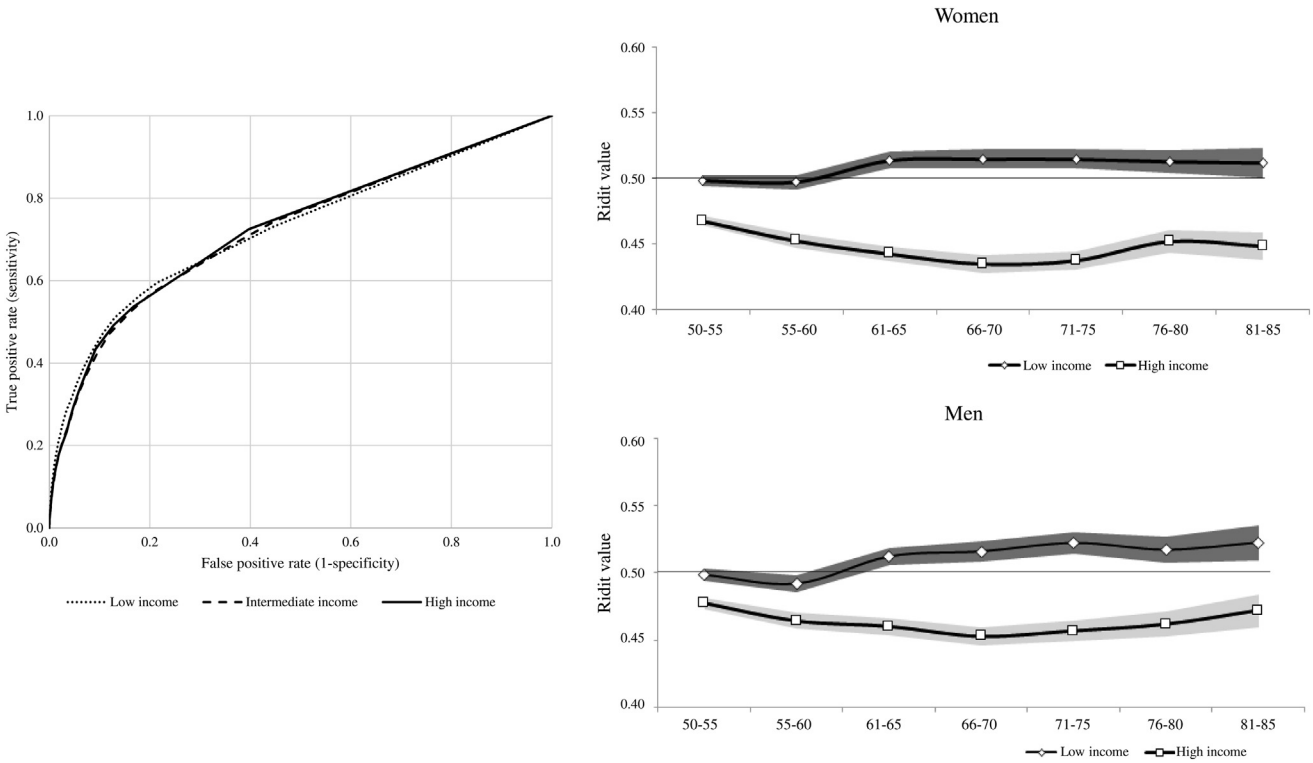

Fig. 5. Receiver Operating Characteristics curves comparing discriminant power of Chronic-Related Score (CReSc) in predicting 5-year mortality according to income category (left box) and age specific values of mean ridit values of the CReSc distributions in women and men according to their income category, with corresponding 95% confidence intervals (right boxes). Mean ridit measures the relative probability that a randomly selected National Health Services (NHS) beneficiary among those belonging to a given income category has a CReSc value indicating better (ridit < 0.5) or worse (ridit > 0.5) health profile than a randomly selected NHS beneficiary among those with intermediate income.

Second, CReSc showed good performance in predicting 5-year mortality among men and women aged between 50 and 85 years, whereas it failed in NHS beneficiaries younger than 50 years and in those older than 85 years. It is not surprising that the few individuals aged 50 years or younger who suffered from chronic diseases mainly experienced death in later life, so the model had poor performance in this age group [33]. On the other hand, the almost total absence of prognostic score specifically developed to predict mortality and other outcomes in elderly people is even not surprising [34], given that almost all people aged 85 years or over suffer from chronic diseases.

Third, consistently with the reported socioeconomic health inequality [35], even observed in countries with universal access to essential health care services such as Italy [36], we observed worse CReSc among NHS beneficiaries who have the lowest income. As low health care accessibility is expected among the most disadvantaged social class, it may be speculated CReSc worse capture health needs among beneficiaries with low socioeconomic status. However, we found impressive stability of CReSc in predicting mortality regardless of the income. This suggests that, at least among NHS beneficiaries of intermediate age, our score reflects comorbidities of the population covered by the NHS, rather than the accessibility to health care services. By accepting this hypothesis, the observed worse health status associated with low and low-to-mild income may aid the identification of vulnerable subgroups who may benefit from tailored health education and management [37,38].

The present study has several strengths. First, because in Italy a public-funded health care system involves virtually all citizens, our sample included all the NHS beneficiaries thus resulting for an unselected population (i.e., ours is a population-based investigation) and very large sample size. Second, CReSc was validated and tested on a random sample of almost 3 million of NHS beneficiaries, a sample so large that random uncertainty slightly affected our estimates. Third, we avoided the selection of comorbidities based on the opinion of experts [39,40] and prevalence data [41,42]. Finally, as the selected diseases were detected through the use of health services, our data overcome the artificial positive socioeconomic inequalities in chronic disease prevalence when using self-reports [43].

Potential limitations must take into account for interpreting or findings. First, our scoring system did not capture health services supplied by private providers. However, due to the universal coverage for essential healthcare, it is unlikely that diseases strongly affecting mortality may escape from the Italian NHS. Second, misdiagnosis (often due to poor accuracy in reporting diagnoses and comorbidities [44]) and upcoding (sometimes in pursuit of higher reimbursements [45]) in hospital records might have generated

too conservative estimates of CReSc performance. However, because diagnostic errors alike affect the compared diagnosis-based comorbidity scores, this issue does not question our main result that CReSc had better performance than both CCI and MCS. Third, since the included outcomes can be considered proxies of the quality of care (i.e., unsuccessful discharge processes or inadequate social care [46] influence mortality and hospital readmissions), our scoring system might not be generalizable to other settings of Italy. Fourth, Lombardy has a long tradition in the management of several health care databases. However, data on outpatient services (including visits and diagnostic tests respectively performed in specialist ambulatories and laboratories accredited by the NHS), payment exemptions, drugs directly administered in the inpatient setting, and emergency room visits are not yet available in several Italian regions. Fifth, generalizability may also be limited by the split approach used in our study since randomly splitting the whole dataset into a training, and a validation set raised concerns from some authors [47]. External validations in other settings (e.g., other Italian regions, countries outside Italy, and different calendar times) should be carried out for ensuring external validity of CReSc. Finally, we must be aware that CReSc may not apply to every relevant outcome and cannot really predict the individual conditions in increasing patients' relative risk of death. For example, our score cannot take into account (1) the conditions that do not affect 5-year mortality, if nor marginally (e.g., type I and type II diabetes), (2) NHS beneficiaries suffering a given condition who did not leave ''footprints'' of routine medical care able to detect that condition (e.g., untreated hypertension) or who mostly escape capture with administrative data (e.g., chronic obstructive pulmonary disease), and (3) patients who did not survive at least 2 years after the onset of an acute condition (e.g., fatal myocardial infarction).

In summary, we developed and validated a prognostic score derived from data usually used for health system management of Lombardy, useful for predicting short-term and long-term mortality, hospitalization, and health costs of each individual NHS beneficiary. CReSc can represent a useful tool for epidemiologists (who need an instrument for risk adjustment in clinical and epidemiological studies), clinicians (who need detecting and managing frail patients in everyday medical practice), and policy-makers (who need assessing health system performance and health policy planning). In particular, we expect that the introduction of the CReSc risk stratification tool, jointly of monetary incentive for services agreeing with proactive management of high-risk patients, may generate clinically and operationally important effects.

CRediT authorship contribution statement

Federico Rea: Methodology, Software, Formal analysis, Writing - review & editing. **Giovanni Corrao:** Conceptualization, Methodology, Writing - original draft, Supervision. **Monica Ludergnani:** Software, Formal analysis, Writing - review & editing. **Luigi Cajazzo:** Conceptualization, Resources, Writing - review & editing. **Luca Merlino:** Conceptualization, Resources, Writing - review & editing.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2019.08.009.

References

[1] World Health Organization. Noncommunicable diseases country profiles 2011. World Health Organization; 2011. Available at https://apps.who.int/iris/handle/10665/44704.
[2] Abegunde DO, Mathers CD, Adam T, Ortegon M, Strong K. The burden and costs of chronic diseases in low-income and middle-income countries. Lancet 2007;370:1929−38.
[3] Lewis G, Kirkham H, Duncan I, Vaithianathan R. How health systems could avert 'triple fail' events that are harmful, are costly, and result in poor patient satisfaction. Health Aff (Millwood) 2013;32:669−76.
[4] Snooks H, Bailey-Jones K, Burge-Jones D, Dale J, Davies J, Evans BA, et al. Effects and costs of implementing predictive risk stratification in primary care: a randomised stepped wedge trial. BMJ Qual Saf 2018;28:697−705.
[5] Snooks H, Bailey-Jones K, Burge-Jones D, Dale J, Davies J, Evans B, et al. Predictive risk stratification model: a randomised stepped-wedge trial in primary care (PRISMATIC). Southampton (UK): NIHR Journals Library; 2018.
[6] Lewis G. Next steps for risk stratification in the NHS. London: NHS England; 2015.
[7] Nuffield Trust. Choosing a predictive risk model: a guide for commissioners in England. London: Nuffield Trust; 2011.
[8] Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. J Clin Epidemiol 1992;45:197−203.
[9] Yurkovich M, Avina-Zubieta JA, Thomas J, Gorenchtein M, Lacaille D. A systematic review identifies valid comorbidity indices derived from administrative health data. J Clin Epidemiol 2015;68:3−14.
[10] Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 1992;45:613−9.
[11] Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. J Clin Epidemiol 1993;46:1075−90.
[12] Romano PS, Roos LL, Jollis JG. Further evidence concerning the use of a clinical comorbidity index with ICD-9-CM administrative data. J Clin Epidemiol 1993;46:1085−90.
[13] Ghali WA, Hall RE, Rosen AK, Ash AS, Moskowitz MA. Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data. J Clin Epidemiol 1996;49:273−8.
[14] Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, 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;173:676−82.
[15] 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;40:373−83.
[16] D'Hoore W, Sicotte C, Tilquin C. Risk adjustment in outcome assessment: the Charlson comorbidity index. Methods Inf Med 1993;32:382−7.
[17] D'Hoore W, Bouckaert A, Tilquin C. Practical considerations on the use of the Charlson comorbidity index with administrative data bases. J Clin Epidemiol 1996;49:1429−33.

[18] Corrao G, Rea F, Di Martino M, De Palma R, Scondotto S, Fusco D, et al. Developing and validating a novel multisource comorbidity score from administrative data: a large population-based cohort study from Italy. BMJ Open 2017;7:e019503.
[19] Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med 1997;16:385–95.
[20] 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;64:749–59.
[21] Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–72.
[22] 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;160:122–31.
[23] Jaja BNR, Saposnik G, Lingsma HF, Macdonald E, Thorpe KE, Mamdani M, et al. SAHIT collaboration. Development and validation of outcome prediction models for aneurysmal subarachnoid haemorrhage: the SAHIT multinational cohort study. BMJ 2018;360:j5745.
[24] Yu W, Xu W, Zhu L. A modified Hosmer–Lemeshow test for large data sets. Commun Stat Theor Methods 2017;46:11813–25.
[25] Bang H, Tsiatis AA. Estimating medical costs with censored data. Biometrika 2000;87:329–43.
[26] Pickett KE, Wilkinson RG. Income inequality and health: a causal review. Soc Sci Med 2015;128:316–26.
[27] Atkinson AB, Rainwater L, Smeeding TM. Income distribution in OECD countries. Paris: OECD; 1995.
[28] Bross IDJ. How to use ridit analysis. Biometrics 1958;3:189–209.
[29] Bross IDJ. Ridt analysis. Am J Epidemiol 1978;107:264.
[30] Mackenbach JP, Kunst AE. Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. Soc Sci Med 1997;44:757–71.
[31] Manor O, Matthews S, Power C. Comparing measures of health inequality. Soc Sci Med 1997;45:761–71.
[32] Torsheim T, Currie C, Boyce W, Kalnins I, Overpeck M, Haugland S. Material deprivation and self-rated health: a multilevel study of adolescents from 22 European and North American countries. Soc Sci Med 2004;59:1–12.
[33] Muller DC, Murphy N, Johansson M, Ferrari P, Tsilidis KK, Boutron-Ruault MC, et al. Modifiable causes of premature death in middle-age in Western Europe: results from the EPIC cohort study. BMC Med 2016;14:87.
[34] Quinzler R, Freitag MH, Wiese B, Beyer M, Brenner H, Dahlhaus A, et al. A novel superior medication-based chronic disease score predicted all-cause mortality in independent geriatric cohorts. J Clin Epidemiol 2019;105:112–24.
[35] Anonymous. Socio-economic inequality in science is on the rise. Nature 2016;537:450.
[36] Alicandro G, Sebastiani G, Bertuccio P, Zengarini N, Costa G, La Vecchia C, et al. The main causes of death contributing to absolute and relative socio-economic inequality in Italy. Public Health 2018;164:39–48.
[37] Kim S, Lee B, Park M, Oh S, Chin HJ, Koo H. Prevalence of chronic disease and its controlled status according to income level. Medicine (Baltimore) 2016;95:e5286.
[38] Cainzos-Achirica M, Capdevila C, Vela E, Cleries M, Bilal U, Garcia-Altes A, et al. Individual income, mortality and healthcare resource use in patients with chronic heart failure living in a universal healthcare system: a population-based study in Catalonia, Spain. Int J Cardiol 2019;277:250–7.
[39] Normand SL, Morris CN, Fung KS, McNeil BJ, Epstein AM. Development and validation of a claims based index for adjusting for risk of mortality: the case of acute myocardial infarction. J Clin Epidemiol 1995;48:229–43.
[40] Desai MM, Bogardus ST Jr, Williams CS, Vitagliano G, Inouye SK. Development and validation of a risk-adjustment index for older patients: the high-risk diagnoses for the elderly scale. J Am Geriatr Soc 2002;50:474–81.
[41] Fleming ST, Pearce KA, McDavid K, Pavlov D. The development and validation of a comorbidity index for prostate cancer among Black men. J Clin Epidemiol 2003;56:1064–75.
[42] Holman CD, Preen DB, Baynham NJ, Finn JC, Semmens JB. A multipurpose comorbidity scoring system performed better than the Charlson index. J Clin Epidemiol 2005;58:1006–14.
[43] Vellakkal S, Millett C, Basu S, Khan Z, Aitsi-Selmi A, Stuckler D, et al. Are estimates of socioeconomic inequalities in chronic disease artefactually narrowed by self-reported measures of prevalence in low-income and middle-income countries? Findings from the WHO-SAGE survey. J Epidemiol Community Health 2015;69:218–25.
[44] Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323–37.
[45] Schonberger RB, Dutton RP, Dai F. Is there evidence for systematic upcoding of ASA physical status coincident with payer incentives? A regression discontinuity analysis of the National Anesthesia Clinical Outcomes Registry. Anesth Analg 2016;122:243–50.
[46] Tsai TC, Joynt KE, Orav EJ, Gawande AA, Jha AK. Variation in surgical-readmission rates and quality of hospital care. N Engl J Med 2013;369:1134–42.
[47] Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol 2016;69:245–7.