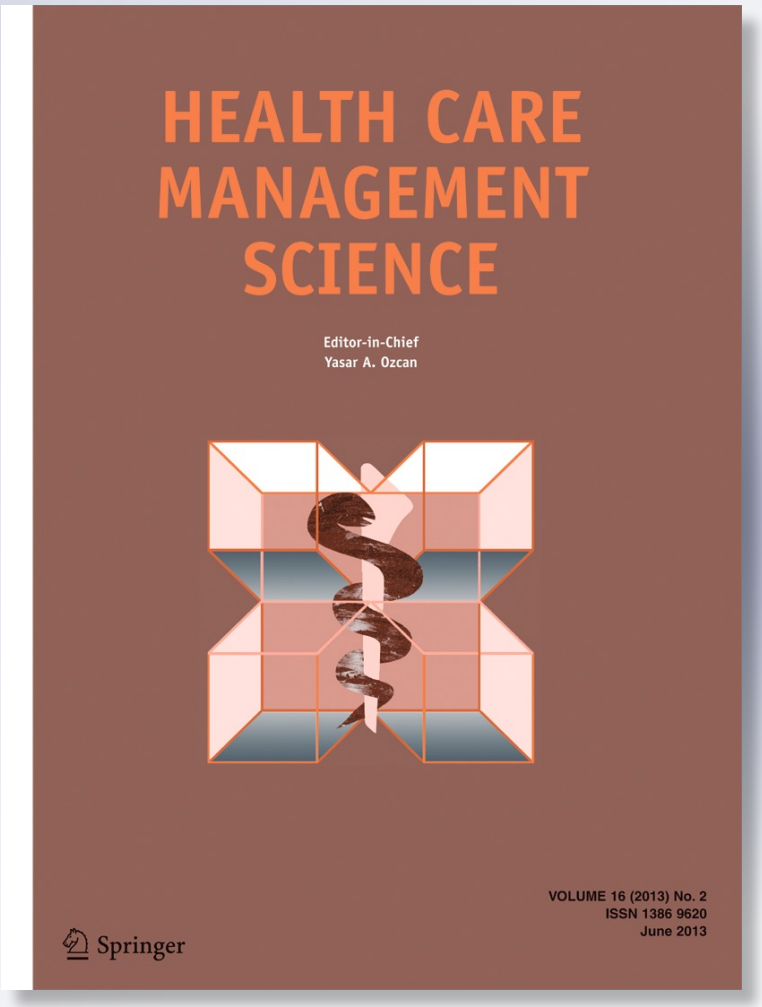
Predicting 30-day all-cause hospital readmissions

Mollie Shulan, Kelly Gao & Crystal Dea Moore

Health Care Management Science

ISSN 1386-9620
Volume 16
Number 2

Health Care Manag Sci (2013)
16:167-175
DOI 10.1007/s10729-013-9220-8



Springer

Health Care Manag Sci (2013) 16:167–175
DOI 10.1007/s10729-013-9220-8

Predicting 30-day all-cause hospital readmissions

Mollie Shulan · Kelly Gao · Crystal Dea Moore

Received: 30 October 2012 / Accepted: 16 January 2013 / Published online: 27 January 2013

Abstract Hospital readmission rate has been broadly accepted as a quality measure and cost driver. However, success in reducing readmissions has been elusive. In the US, almost 20 % of Medicare inpatients are rehospitalized within 30 days, which amounts to a cost of $17 billion. Given the skyrocketing healthcare cost, policymakers, researchers and payers are focusing more than ever on readmission reduction. Both hospital comparison of readmissions as a quality measure and identification of high-risk patients for post-discharge interventions require accurate predictive modeling. However, most predictive models for readmissions perform poorly. In this study, we endeavored to explore the full potentials of predictive models for readmissions and to assess the predictive power of different independent variables. Our model reached the highest predicting ability (c-statistic =0.80) among all published studies that used administrative data. Our analyses reveal that demographics, socioeconomic variables, prior utilization and Diagnosis-related Group (DRG) all have limited predictive power; more sophisticated patient stratification algorithm or risk adjuster is desired for more accurate readmission predictions.

Keywords Hospital readmissions · Logistic regression · Predictive power

MSC: 62P10

M. Shulan (✉) · K. Gao
Department of Veterans Affairs, Stratton VA Medical Center,
113 Holland Avenue,
Albany, NY 12208, USA
e-mail: mollie.shulan@va.gov

K. Gao · C. D. Moore
Department of Social Work, Skidmore College, 815 N. Broadway,
Saratoga Springs, NY 12866, USA

1 Introduction

With ever-rising healthcare costs in many countries around the world, policymakers, researchers and payers are increasingly focusing on reduction of hospital readmissions [1–7]. In comparison to other countries, the US appears to have the highest readmission rate although varied by conditions [5, 6, 8]. Studies showed almost one-fifth of Medicare inpatients are rehospitalized within 30 days, which translates into a cost of $17 billion or nearly 20 % of Medicare's total hospital payment [1]. The Medicare Payment Advisory Commission (MedPAC) estimated Medicare spends $12 billion annually for readmissions deemed potentially preventable [9]. The Centers for Medicare & Medicaid Services (CMS) in 2009 began publicly reporting readmission rates for certain conditions, and now has started since October 1, 2012 to reduce payments to the hospitals with excessive readmissions for acute myocardial infarction (AMI), heart failure, and pneumonia [10, 11]. Furthermore, CMS is now considering expanding the list beyond the three conditions covered by the program in the near future [4].

Despite the necessity of reducing hospital readmissions, broad intervention for all discharged patients is not feasible due to resource constraints; programs aimed to reduce readmissions are more sustainable if high-risk patients can be accurately identified and targeted [11, 12]. In addition to post-discharge intervention, readmission predictive models can also be used for hospital comparison of readmissions as a measure of quality of care. However, none of these applications would be adequate if predicting or discriminative ability is low [11, 12]. According to a recent review study, most current risk prediction models for readmissions perform poorly (c-statistic<0.7); a few are acceptable (c-statistic= 0.72), and only one study reached good prediction accuracy (c-statistic=0.83) but the model was based on patient survey data that were not routinely available in most hospitals or health insurance plans [11, 13].

With constantly advancing information technology, administrative data are becoming increasingly closer to real time and offering greater potentials in improving quality of care as well as containing costs. In this study, we sought to explore the full capacity of administrative data in predicting all-cause readmissions. The aim of our study is two-fold: (1) to explore the maximum discriminative ability of readmission predictive models using administrative data, and (2) to assess the predictive power of different independent variables, i.e., demographic, socioeconomic, prior utilization, patient risk, comorbidities and their interactions.

2 Data source and study variables

In this study, we analyzed inpatient data from Veterans Healthcare Network Upstate New York (VISN 2), which is one of the 21 Networks through which the US Department of Veterans Affairs (VA) delivers care to its 5.6 million patients annually. VISN 2, with five medical centers and 31 outpatient clinics across upstate New York, serves approximately 140,000 patients with an annual budget of one billion dollars. In fiscal year (FY) 2011, 8,718 patients were hospitalized for acute care in four of the medical centers in VISN 2 (one medical center does not have inpatient services).

We analyzed all of the 8,718 patients for readmissions. We define a hospital readmission as a hospitalization within 30 days after the initial or index discharge. As in other studies, no more than one readmission was counted in this study [11]. Patients transferred from other hospitals were excluded. To count readmission accurately, we also used the first month of FY 2012 (October 2011) data to capture readmissions from the index hospitalizations in the last month of FY 2011 (September, 2011).

VA National Patient Care Database (NPCD) hosted at the Austin Information Technology Center (AITC) was the primary data source for this study. We used Patient Treatment File (PTF), associated Census File in FY2011 and first 30 days of FY2012 PTF to identify index hospitalizations and readmissions. In addition to encounter data such as admission/discharge dates and ICD-9 CM codes, PTF also contains patient demographic and socioeconomic variables such as age, gender, race and income. NPCD including PTF is the gold standard for VA operational analysis and research. Most of the data fields such as admission/discharge dates and clinical information like ICD-9 CM codes are routinely and rigorously validated with strict business rules. Its income information is means tested. One exception is that its race information is often incomplete because VA does not mandate veterans to report race status. However, for the last several years, VA has systematically gathered race information from other data sources such as Medicare and Department of Defense (DOD); as a result, the updated race data is deemed accurate and reliable [14, 15].

We also used Decision Support System (DSS) files that contain actual patient care costs rather than claims or paid as in private health plans. DSS costs are the primary financial data for internal operations and congressional inquiries. For case-mix, VA has been using DxCG, commercially available software [16], to systematically measure risk and comorbidities of all 5.6 million patients for the last decade. DxCG algorithm is solely based on administrative data: it uses ICD-9 CM codes, age and gender as input data to classify patients into hierarchical condition category (HCCs, which are similar to the HCCs used by CMS) and then produces a risk score for each patient by applying cost-weight from commercial or Medicare population [16–20].

Based on the literature and data availability in the VA, we collected and used the independent variables for this study in four categories: (1) demographics: age, sex, marital status, number of dependents, race; (2) socioeconomic variables: patient income, and patient insurance status, i.e. not covered by any insurance (equals 1, otherwise 0), enrolled in Medicare (equals 1, otherwise 0), enrolled in Medicaid (equals 1, otherwise 0), and covered by private insurance (equals 1, otherwise 0); (3) prior utilization and cost: length of stay (LOS) of the index hospitalization, prior year (FY2010) number of primary care visits, prior year number of emergency department (ED) visits and the prior year cost; and (4) risk and comorbidities: Diagnosis-related Group (DRG), DxCG and its HCCs. Notice that case-mix or risk adjustment has been an imperfect science and selection/use of risk and comorbidities in the literature has been rather ad hoc. Some studies only used a priori coexisting conditions ranging from a few to a couple of dozen [1, 2, 21, 22] while others used a comprehensive measure (e.g., Charlson comorbidity index, Medicare mortality prediction system score) and a set of coexisting conditions [23, 24]. In this study, we chose DRG, DxCG and its HCCs as the patient risk and comorbidity measures due to data reliability and availability. DxCG is a well validated risk adjustor; most studies found DxCG was superior to other algorithms in predicting resources use [17–20]. In this study, both DRGs and HCCs were used in the regressions as dummy variables, i.e. equals 1 if a patient is in a DRG or HCC otherwise equals 0.

This study used no identifiable patient private information and therefore exempted from IRB review under VA Title 38, Section 16.101(b)(4).

3 Modeling and analysis

We employed logistic regression to predict the probability of rehospitalizations. Logistic regression has been the most

extensively used model in predicting readmissions or other outcomes where the dependent variable is binary, i.e., equals 1 if the event happened, otherwise equals 0. The model's predicting or discriminative ability is measured by the c-statistic, which is defined as the proportion of times the model correctly discriminates a random pair of individuals with or without the outcome. It is also equivalent to the area under the receiver operating characteristic curve. A c-statistic of 0.5 means that the model is no better than a random pick; a c-statistic of 0.7–0.8 suggests that the model has modest discriminative ability; and a c-statistic of 0.8 or greater suggests good discriminative ability [11]. To prevent model over-fit, we only included variables with p-values less than 0.1 (by stepwise) in the final regression analysis, and we also calculated shrinkage coefficient, an indicator of over-fit [25]. To be prudent, we further validated the model by split-sample method [25, 26], which is also referred to as 2-fold cross-validation. With this method, the full sample was randomly split into a derivation sample (50 %) and a validation sample (50 %). The model was estimated on the derivation sample and then the estimated coefficients were applied to the validation sample. The analyses were conducted by using PROC LOGISTIC of SAS 9.3.

To demonstrate the predictive power of different independent variables, we configured and conducted six models from basic to comprehensive. Model 1: only demographic, socioeconomic variables and fixed effects of the medical centers are included in the regression as the independent variables, i.e., age, sex, marital status, race, income, number of dependents, enrolled in Medicare, enrolled in Medicaid or covered by other private insurance (no insurance status was omitted in the regression as reference). We used three dummy variables (one is omitted as reference) as the fixed effect to control for potentially different practice patterns among the four medical centers. Model 2: variables in model 1, and prior year utilizations, i.e., prior year's primary care visits, ED visits, cost, and the LOS of initial hospitalization. Model 3: variables in model 2, and DRGs (a dummy variable is created for each DRG). Model 4: variables in model 3, and the patient DxCG risk score. Model 5: variables in model 4, and comorbidities measured by HCCs. Model 6: variables in model 5, and the major interactions of the independent variables. We tested the interactions of each variable of the demographic, social economic, prior cost/utilization with case-mix (DxCG, DRGs and HCCs); from HCCs we also picked a priori of major conditions (congestive heart failure, chronic obstructive pulmonary disease, diabetes, hypertension, hyperlipidemia, depression, obesity, and smoking) to interact with all other conditions. All quadratic terms of the variables (except for the dummy variables) were tested too, but inclusion or exclusion of the variables in the final regression depends on the p-values in the stepwise procedure.

4 Results

The descriptives of the independent variables are reported in Table 1. Without loss of much information, DRGs and HCCs that are dummy variables (106 and 271 dummy variables for DRGs and HCCs respectively in this study population), the interactions and quadratic terms are not reported in Table 1. Also, only the parameter estimates of the final full model are reported in Table 2.

In the first model that only included demographic/socioeconomic characteristics and the fixed effect of the medical centers, 4 variables were statistically significant (p-values< 0.05) and kept in the model. The c-statistic was 0.55 (95 % CI: 0.54–0.57). In the second model, seven variables were statistically significant (six variables with p-values <0.05, one with p-value <0.1) and kept in the model. The c-statistic was 0.60 (95 % CI: 0.59–0.62). In the third model, 21 variables met the inclusion criterion and were kept in the model (14 variables with p-values <0.05, others <0.1), and the c-statistic was 0.63 (95 % CI: 0.61–0.64). In the fourth model, 22 variables met the inclusion criterion and were kept in the model (16 variables with p-values <0.05, others <0.1), and the c-statistic was 0.73 (95 % CI: 0.71–0.74). In the fifth model, 74 variables met the inclusion criterion and were kept in the model (63 variables with p-values <0.05, others <0.1), and c-statistic was 0.78 (95 % CI: 0.77–0.80). As reported in Table 2, the final full model includes 93 variables (87 variables with p-values <0.05, others <0.1) and c-statistic reached 0.80 (95 % CI: 0.79–0.81). The receiver operating characteristic curves of all six models are reported in Fig. 1.

The full model in this study had an estimated shrinkage coefficient of 0.94 which indicates no over-fit [25]. Nevertheless, we further validated our model by using a data splitting technique (the data was randomly selected into two equal size samples): for the derivation sample, c-statistic is 0.80 (95 % CI: 0.79–0.82); for the validation sample c-statistic is 0.79 (95 % CI: 0.78–0.81). In order to visually further examine the prediction accuracy, we compared the predicted and observed readmissions by five estimated risk categories. As can be seen from Fig. 2, in the derivation sample, the predicted and the observed were very close; in the validation sample, the model slightly overestimated the number of readmissions for the low risk group and slightly underestimated the number of readmissions for the high risk group.

5 Discussion

Public reporting and financial penalties were intended to compel hospitals with high readmission rates to improve quality and reduce cost; however, studies have concluded

Table 1 Descriptives of Independent Variables

Variables	Patients without Readmission (n=7,310) Mean or Percent (SD)	Patients with Readmission (n=1,408) Mean or Percent (SD)	P-value
Age	66.43 (15.37)	67.04 (14.65)	0.170
Sex (male)	0.94 (0.24)	0.96 (0.20)	0.005
Marital Status (married)	0.40 (0.49)	0.37 (0.48)	0.044
Number of Dependents	0.21 (0.40)	0.2 (0.40)	0.470
Race Status (Black)	0.11 (0.31)	0.12 (0.32)	0.507
Patient Income	25,202 (41,380)	24,631 (38,658)	0.632
No Health Insurance	0.27 (0.44)	0.22 (0.42)	<0.001
Enrolled in Medicare	0.60 (0.49)	0.66 (0.47)	<0.001
Enrolled in Medicaid	0.02 (0.13)	0.03 (0.17)	0.002
Covered by Private Insurance	0.11 (0.31)	0.08 (0.28)	0.006
Index Hospitalization Length of Stay	5.45 (7.93)	6.19 (8.96)	0.002
Prior Year Primary Care Visits	3.50 (3.90)	3.9 (3.64)	<0.001
Prior Year Emergency Department Visits	1.18 (2.07)	2.02 (3.29)	<0.001
Prior Year Patient Cost (in logarithm)	8.65 (2.85)	9.33 (2.54)	<0.001
DxCG Risk Score	2.85 (2.40)	4.59 (3.25)	<0.001
Medical Center A	0.20 (0.40)	0.19 (0.40)	0.577
Medical Center B	0.05 (0.22)	0.04 (0.20)	0.080
Medical Center C	0.35 (0.48)	0.35 (0.48)	0.635
Medical Center D	0.40 (0.49)	0.41 (0.49)	0.446

that unless risk prediction and risk adjustment of readmissions become more accurate, it is inappropriate to compare hospitals and reimburse them accordingly [11, 27, 28]. Even more importantly, the feasibility of comprehensive post-discharge interventions aimed to reduce readmissions hinges on the accuracy of the predictive models due to limited healthcare resources. However, a recent systematic literature review revealed that "most current readmission risk prediction models that were designed for either comparative or clinical purposes perform poorly…efforts to improve their performance are needed as use becomes more widespread" [11]. Among the 30 studies reviewed, one reached a c-statistic of 0.83 by using patient survey data, two reached a c-statistic of 0.72 and all others had c-statistics less than 0.70. Even the model developed by CMS for hospital profiling can only reach c-statistics of 0.56–0.66 [11, 12].

In this study, we sought to improve the model discriminative ability and to assess the predictive power of different independent variables. We purposely restricted our model to administrative data that are routinely available to all health plans and government agencies; laboratory test results and vital signs were not used because of their lack of availability in other healthcare systems. Based on administrative data, our model reached higher discriminative ability (c-statistic= 0.80) than any other published studies using administrative data. To reassure the high c-statistic was not due to over fitting, in addition to the 50 % split of the study population for model validation, we also split the study population by 75 % (derivation sample) and 25 % (validation sample), which gave almost identical results.

In research or practice, one can rarely see two models using different data result in the same variables being statistically significant; thus, "copying" predictive models from one healthcare system to another could easily result in misspecification. Even within the same healthcare system, predictive models should be reconfigured over time (e.g. test different function forms, add new variables or drop those becoming statistically insignificant). To add insight into predictive model development in different health care systems, we assessed and reported the predictive powers of different independent variables. Our findings indicate that demographics only had a predicting power of 5 % better than pure chance, prior utilization/cost added another 5 %, and DRGs can only improve less than 3 % on top of that, which collectively yielded a c-statistic of 0.63. The strongest predicting power came from the comprehensive risk score measured by DxCG, which improved the model by 10 %, HCCs added another 5 %, but the interactions only increased predictive ability by less than 2 % to a final c-statistic of 0.80.

Due to the use of stepwise method with a p-value of 0.1, apart from the comorbidities (DRGs and HCCs), therefore only a few "stand alone" variables left in the final regression. The statistical significance of the quadratic term of DxCG indicated the effect of DxCG on readmissions was not linear. Among all the prior cost and utilization variables

Table 2 Logistic Regression Parameter Estimate, Odds Ratio and Confidence Interval

Variable Category	Independent Variable	Parameter Estimate	P-value	Odds-Ratio	Confidence Interval (95 %)
Demographics	Sex (male)	0.632	0.001	1.88	1.31 – 2.71
Prior Utilization	Index Hospitalization Length of Stay	−0.014	0.001	0.99	0.98 – 0.99
DRG	Chronic Obstructive Pulmonary Disease w MCC	0.815	0.011	2.26	1.21 – 4.23
DRG	Chronic Obstructive Pulmonary Disease w/o CC/MCC	1.062	<0.001	2.89	1.88 – 4.45
DRG	Circulatory Disorders Except AMI, W Card Cath w/o MCC	0.582	0.033	1.79	1.05 – 3.06
DRG	Major Joint Replacement or Reattachment of Lower Extremity w/o MCC	−1.370	0.011	0.25	0.09 – 0.73
DRG	Signs & symptoms of Musculoskeletal System & Conn Issue w/o MCC	1.121	0.005	3.07	1.41 – 6.69
DRG	Septicemia or Severe Sepsis w/o MV 96+ Hours w/o MCC	−1.227	0.059	0.29	0.08 – 1.05
DRG	Psychoses	0.496	0.001	1.64	1.22 – 2.22
DxCG Score	DxCG Risk Score Squared	−0.013	<0.001	0.99	0.98 – 0.99
DxCG Score	DxCG Risk Score	0.372	<0.001	1.45	1.36 – 1.55
HCC	Secondary Cancers Except Lymph Node / Disseminated Cancer	−0.712	<0.001	0.49	0.35 – 0.68
HCC	Melanoma	1.016	0.024	2.76	1.14 – 6.68
HCC	Disorders of Fluid/Electrolyte/Acid–base Balance	0.446	<0.001	1.56	1.35 – 1.81
HCC	Other Endocrine/Metabolic/Nutritional Disorders	0.712	0.011	2.04	1.17 – 3.54
HCC	Acute Pancreatitis and Other Pancreatic Disease	0.470	0.011	1.60	1.11 – 2.30
HCC	Polyarteritis Nodosa, Systemic Lupus Erythematosus Related Conditions	0.679	0.034	1.97	1.05 – 3.69
HCC	Withdrawal and Other Specified Drug-Induced Mental Disorders	0.742	0.001	2.10	1.36 – 3.24
HCC	Alcohol Psychosis	0.657	<0.001	1.93	1.42 – 2.61
HCC	Schizoid and Borderline Personality Disorders	0.586	0.020	1.80	1.10 – 2.95
HCC	Cardiac Arrest / Ventricular Fibrillation/Flutter	−0.666	0.027	0.51	0.28 – 0.93
HCC	Acute Heart Failure and CHF Exacerbation	0.537	0.001	1.71	1.26 – 2.33
HCC	Specified Heart Arrhythmias	0.408	<0.001	1.50	1.27 – 1.78
HCC	Deep Vein Phlebitis/Thrombosis and Embolism, Leg	0.700	<0.001	2.01	1.37 – 2.97
HCC	Pneumococcal Pneumonia, Empyema, Lung Abscess	0.775	0.001	2.17	1.36 – 3.47
HCC	Viral and Unspecified Pneumonia, Pleurisy	0.247	0.008	1.28	1.07 – 1.54
HCC	Asthma	0.579	0.004	1.78	1.21 – 2.64
HCC	Urinary Calculi / Stricture / Reflux / Retention	0.360	<0.001	1.43	1.18 – 1.74
HCC	Acute and Other Urinary Tract Infection	0.290	0.002	1.34	1.11 – 1.60
HCC	Perimenopausal, Other Congenital, Acquired Female Genital Disorders	0.874	0.030	2.40	1.09 – 5.27
HCC	Cellulitis, Local Skin Infection	0.356	<0.001	1.43	1.20 – 1.70
HCC	Poisonings and Allergic Reactions	0.701	<0.001	2.01	1.67 – 2.43
HCC	Other Major Symptoms, Abnormalities	0.336	<0.001	1.40	1.22 – 1.60
HCC	Screening/Observation/Special Exams	1.299	0.005	3.66	1.47 – 9.12
HCC	Other Postsurgical States / Elective Surgery / Aftercare	1.281	<0.001	3.60	2.35 – 5.52
Interactions	Age*Nutritional Deficiency and Other/Unspecified Anemias	0.003	0.008	1.00	1.00 – 1.01
Interactions	Age*Drug Induced Hallucinations, Delusions, and Delerium	−0.022	0.050	0.98	0.96 – 1.00
Interactions	Age*Alcohol Dependence	0.004	0.022	1.00	1.00 – 1.01
Interactions	Age*Bipolar Disorder	0.011	<0.001	1.01	1.01 – 1.02
Interactions	Age*Prolonged Posttraumatic Stress Disorder	−0.003	0.066	1.00	0.99 – 1.00
Interactions	Age*Attempted Suicide / Self-Inflicted Injury	0.007	0.066	1.01	1.00 – 1.01
Interactions	Age*Motor Neuron Disease and Spinal Muscular Atrophy	0.013	0.040	1.01	1.00 – 1.03
Interactions	Age*Respirator Dependence/Tracheostomy Status	−0.016	0.003	0.98	0.97 – 0.99
Interactions	Age*Unstable Angina and Other Acute Ischemic Heart Disease	0.007	<0.001	1.01	1.00 – 1.01
Interactions	Age*Hypotension, Excluding Hypotension of Hemodialysis	0.005	<0.001	1.00	1.00 – 1.01
Interactions	Age*Decubitus Ulcer of Skin	−0.006	0.005	0.99	0.99 – 1.00
Interactions	Age*Other Postsurgical States / Elective Surgery / Aftercare	−0.013	<0.001	0.99	0.98 – 0.99
Interactions	Age*Infection/Inflammation From Internal Device/Implant/Graft	0.008	0.011	1.01	1.00 – 1.01

Table 2 (continued)

Variable Category	Independent Variable	Parameter Estimate	P-value	Odds-Ratio	Confidence Interval (95%)
Interactions	Age*Central Nervous System Infection	0.024	<0.001	1.02	1.01 – 1.04
Interactions	Age*Other and NOS Soft Tissue Disorders	0.005	0.012	1.00	1.00 – 1.01
Interactions	Obesity*Schizophrenia	1.497	0.001	4.47	1.84 – 10.86
Interactions	Obesity*Bipolar Disorder	−1.548	0.037	0.21	0.05 – 0.91
Interactions	Obesity*Major Depression	−3.109	0.004	0.04	0.01 – 0.38
Interactions	Obesity*Hypotension, Excluding Hypotension of Hemodialysis	−0.989	0.028	0.37	0.15 – 0.90
Interactions	Obesity*Fistulae, Cysts, Other Urinary System Anomalies	1.594	0.033	4.92	1.14 – 21.29
Interactions	Obesity*Dyspnea, Distress, Other Respiratory/Chest Symptoms	−1.005	0.028	0.37	0.15 – 0.90
Interactions	Obesity*Infection/Inflammation From Internal Device/Implant/Graft	1.844	0.024	6.32	1.28 – 31.32
Interactions	Obesity*Other and NOS Soft Tissue Disorders	1.539	0.003	4.66	1.68 – 12.92
Interactions	Tobacco use*Schizophrenia	0.769	<0.001	2.16	1.56 – 2.99
Interactions	Tobacco use*Depression	0.396	0.006	1.49	1.12 – 1.97
Interactions	Tobacco use*Colorectal Cancers	1.151	0.003	3.16	1.46 – 6.83
Interactions	Tobacco use*Glaucoma	−0.657	0.003	0.52	0.34 – 0.80
Interactions	Tobacco use*Protein-Calorie Malnutrition	−0.891	0.010	0.41	0.21 – 0.81
Interactions	Tobacco use*Specific Bacterial Infection In Other Diseases	0.655	0.003	1.93	1.24 – 2.98
Interactions	Tobacco use*Ulcer, Perforation, Obstruction, and Peritonitis	0.848	<0.001	2.33	1.45 – 3.76
Interactions	Tobacco use*Fractures of the Knee, Leg, and Multiple Limbs	1.090	0.015	2.97	1.23 – 7.17
Interactions	Hypertension*Anemia in Chronic Illness	0.762	0.001	2.14	1.38 – 3.33
Interactions	Hypertension*Heart Valve Replacement / Transplant	1.058	0.064	2.88	0.94 – 8.83
Interactions	Hypertension*Specified Heart Arrhythmias	−0.500	0.030	0.61	0.39 – 0.95
Interactions	Hypertension*Gangrene/Pulmonary Embolism/Vascular Complications	1.102	0.002	3.01	1.51 – 5.99
Interactions	Hypertension*Other Lung/Laryngeal Anomalies / Oxygen Dependence	1.317	0.029	3.73	1.14 – 12.19
Interactions	Hypertension*Thyroid/Testicular/Ovarian and Metabolic Disorders	−0.498	0.010	0.61	0.42 – 0.89
Interactions	Hypertension*Chronic Pancreatitis and Intestinal Malabsorption	1.465	0.037	4.33	1.09 – 17.1
Interactions	Hypertension*Knee Joint Replacement Status	0.893	0.022	2.44	1.14 – 5.25
Interactions	COPD*Septicemia/Shock	−0.411	0.035	0.66	0.45 – 0.97
Interactions	COPD*Sprains of the Shoulder and Upper Arm	0.896	0.085	2.45	0.89 – 6.78
Interactions	COPD*Disorders of Immunity	−1.027	0.015	0.36	0.16 – 0.82
Interactions	COPD*Dementia	−0.401	0.039	0.67	0.46 – 0.98
Interactions	COPD*Bipolar Disorder	−0.706	0.020	0.49	0.27 – 0.90
Interactions	COPD*Acute Lung Edema / Asphyxia	0.421	0.015	1.52	1.08 – 2.14
Interactions	COPD*Major Eye Surgery	0.490	0.003	1.63	1.18 – 2.25
Interactions	COPD*Cataract	0.388	0.007	1.47	1.11 – 1.96
Interactions	COPD*Fistulae, Cysts, Other Urinary System Anomalies	0.841	0.011	2.32	1.21 – 4.44
Interactions	COPD*Internal Injuries	1.533	0.004	4.63	1.62 – 13.25
Interactions	COPD*Artificial Openings for Feeding or Elimination	0.822	0.007	2.27	1.25 – 4.15
Interactions	COPD*Other Postsurgical States / Elective Surgery / Aftercare	−0.326	0.007	0.72	0.57 – 0.91
Interactions	COPD*Hyperlipidemia	0.288	0.011	1.33	1.07 – 1.67
Interactions	COPD*End-Stage Liver Disease	1.279	0.005	3.59	1.48 – 8.71
Interactions	COPD*Specific Bacterial Infection In Other Diseases	0.526	0.014	1.69	1.11 – 2.57
Interactions	COPD*Enteritis, Hernia, Other Specified Gastrointestinal Disorders	0.404	0.026	1.50	1.05 – 2.14
Interactions	COPD*Hip Fracture/Dislocation	−1.096	0.030	0.33	0.12 – 0.90
Interactions	Major Depression*Surgical and Procedural Infection, Hemorrhage	0.836	0.005	2.31	1.28 – 4.16
Interactions	Major Depression*Other Hepatitis and Chronic Liver Disease	1.928	0.026	6.87	1.25 – 37.69

only LOS was kept in the final model, which indicates shorter LOS was associated with high risk of readmission. Interestingly, it has been often concluded that "the risk of readmission rises with increasing length of stay" [29].

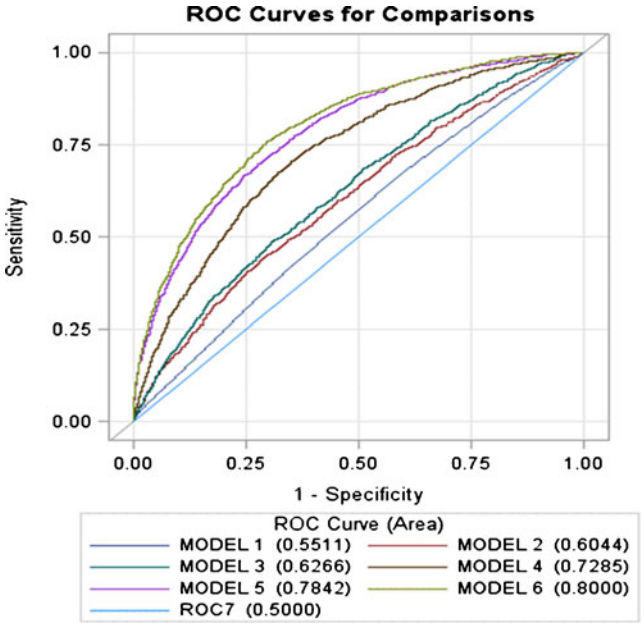


Fig. 1 Receiver operating characteristic curves for 30-day readmissions

However, this relationship is more likely a symptom of lacking sufficient risk adjustment: sicker patients require longer LOS and are more susceptible to rehospitalization. In fact, the univariate analysis in Table 1 shows this specious relationship (p=0.0017), which was reversed after the comprehensive risk adjustment in the full model. Our findings also indicate that male patients were more likely to be rehospitalized compared with female patients, but it has little generalizability because the veteran population in this study only comprises 5 % female patients.

Since the model was configured to focus on prediction, some parameter estimates of the comorbidities may be difficult to explain. This is because the interpretation of coefficients of the dummy variable rests on the reference group that is omitted in the regression. However, with the stepwise rule, many variables (DRGs and HCCs) are dropped off, which are collectively treated as the reference group by the model. As a result, the interpretation of the coefficients (e.g. a negative/positive coefficient means lower/higher risk of readmission) is meaningful only when compared with the composite reference group that is collectively excluded from the model. Another complication is that many of these comorbidities could be highly correlated, which makes the parameter estimates entangled and difficult to interpret. Multicolinearity is a serious and difficult problem if the modeling objective is to assess the effect of each these comorbidities on readmissions, however, it is barely an issue at all if the modeling objective is prediction [30].

Of special note, a p-value of 0.05 is universally accepted as the "magic" number in clinical trials; however, in other empirical studies there is little consensus on the p-value used to include or exclude variables in regressions. Most model selection methods such as information criteria which are based on sum of square errors are designed for continuous dependent variables. Little can be found on variable selection methods for models with binary dependent variables. In the literature, some studies keep all their variables in their final models while others exclude the insignificant ones [1, 24]. It is well known that keeping unlimited number of insignificant variables in the model could easily result in over-fit while omitting statistically significant variables could lose predicting power. In this study, we examined a model including all the variables (except for the interactions) and reached a c-statistic of 0.91 (95 % CI: 0.90–0.92), which appears to be very high. However, the same model can only produce a c-statistic of 0.57 (95 % CI: 0.54–0.60) on the validation sample (50 % split). As described in the method section, we used stepwise method to exclude the insignificant variables; to achieve the highest predictive power, we tried p-values of 0.05, 0.1, 0.15, 0.20 and 0.25. We found that a p-value of 0.2 for a variable to enter and 0.10 to stay in the regression yielded satisfactory results: the highest c-statistics for both

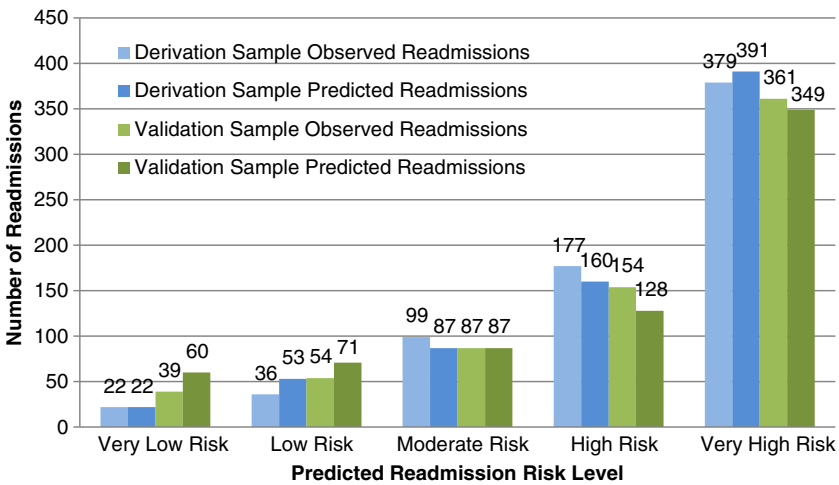


Fig. 2 Comparison of predicted and observed readmissions

derivation and validation samples with the smallest difference between the two c-statistics.

Several limitations to our study need to be noted. First, since VA has routinely used DxCG for case-mix over a decade, for straightforward implementation in the VA system, we also adopted DxCG and its HCCs to adjust patient risk or comorbidities in this study. Although studies found that DxCG is superior in predicting resource use [17–20], we have not identified any studies that compare the predictive power of DxCG on readmissions with other algorithms such as Clinical Risk Groups (CRG), Adjusted Clinical Groups (ACG), and Episode Treatment Groups (ETG), which warrants further research in the future. In addition, publically available groupers such as AHRQ's Clinical Classification Software (CCS) and CMS's HCC software are also worth exploring [31, 32]. Second, the data used in this study were from one region (Upstate New York), a veteran population (e.g. fewer female and more mental health patients; no data on education level and employment status), and a care delivery system with global budget and few financial barriers to access of care; therefore, the results may not be replicable in other health care systems; further, the data are not real time, and as a result, cannot be used as a tool for pre-discharge planning that could have significant effect on readmissions [33, 34]. Third, even though our model can be used in both hospital comparisons and patient identification for readmission reduction, caution needs to be exercised. High predictive power is a necessary but not a sufficient condition for valid comparisons; unlike predictions, hospital comparisons demand more delicate use of confounders (e.g. place of services and health literacy) and more complete data (e.g. Medicare data if implemented in VA). For predictions, the model should not and cannot replace clinical judgment. Finally, our selection of p-values in the stepwise method was rather ad hoc than systematic, and therefore further research is needed to assess the theoretical and empirical relationship between the p-values and c-statistics.

Taken together, we hope this study would render unique insight to the potentials of administrative data in predicting hospital readmissions for enhancing quality of patient care and reducing costs.

Acknowledgments This material is based upon work supported in part by the Office of Research and Development, Department of Veterans Affairs. The authors wish to thank an anonymous statistician for her thorough statistical support.

References

1. Jencks SF, Williams MV, Coleman EA (2009) Rehospitalizations among patients in the Medicare fee-for service program. N Engl J Med 360:1418–1428
2. Jha AK, Orav EJ, Epstein AM (2009) Public reporting of discharge planning and rates of readmissions. N Engl J Med 361:2637–2645
3. Hernandez AF, Greiner MA, Fonarow GC, Hammill BG, Heidenreich PA, Yancy CW, Peterson ED, Curtis LH (2010) Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. JAMA 303(17):1716–1722
4. Pizer SD (2013) Should hospital readmissions be reduced through payment penalties? Med Care 51(1):20–22
5. Monette M (2012) Hospital readmission rates under the microscope. Can Med Assoc 184(12):E651–E652
6. Westert GP, Lagoe RJ, Keskimäki I, Leyland A, Murphy M (2002) An international study of hospital readmissions and related utilization in Europe and the USA. Health Policy 61(3):269–278
7. Snow V, Beck D, Budnitz T, Miller DC, Potter J, Wears RL, Weiss KB, Williams MV (2009) Transitions of care consensus policy statement: American College of Physicians, Society of General Internal Medicine, Society of Hospital Medicine, American Geriatrics Society, American College of Emergency Physicians, and Society for Academic Emergency Medicine. J Gen Intern Med 24(8):971–976
8. Canadian Institute for Health Information (2012) All-cause readmission to acute care and return to the emergence department https://secure.cihi.ca/estore/productFamily.htm?locale=en&pf=PFC1823. Accessed 11 January 2013
9. Medicare Payment Advisory Commission (2008) Report to the congress: reforming the delivery system. MedPAC, Washington
10. Averill RF, McCullough EC, Hughes JS, Goldfield NI, Vertrees JC, Fuller RL (2009) Redesigning the Medicare inpatient PPS to reduce payments to hospitals with high readmission rates. Health Care Financing Rev 30(4):1–15
11. Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, Kripalani S (2011) Risk prediction models for hospital readmission, a systematic review. JAMA 306(15):1688–1698
12. Amarasingham R, Moore BJ, Tabak YP et al (2010) An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data. Med Care 48(11):981–988
13. Coleman EA, Min SJ, Chomiak A, Kramer AM (2004) Post hospital care transitions: patterns, complications, and risk identification. Heal Serv Res 39(5):1449–1465
14. Stroupe KT, Tarlov E, Zhang Q, Haywood T, Owens A, Hynes DM (2010) Use of Medicare and DoD data for improving VA race data quality. J Rehabil Res Dev 47(8):781–795
15. Trivedi AN, Crebia RC, Wright SM, Washinton DL (2011) Despite improved quality of care in the veterans affairs health system, racial disparity persists for important clinical outcomes. Heal Aff 4:707–715
16. DxCG. http://www.veriskhealth.com/content/verisk-health-sightlines-dxcg-risk-solutions?gclid. Accessed 11 January 2013
17. Ellis RP, Ash A (1995) Refinements to the diagnostic cost group (DCG) model. Inquiry 32:418–429
18. Liu CF, Sales AE, Sharp ND, Fishman P, Sloan KL, Todd-Stenberg J, Nichol WP, Rosen AK, Loveland S (2003) Case-mix adjusting performance measures in a veteran population: pharmacy- and diagnosis-based approaches. Heal Serv Res 38(5):1319–1338
19. Sales AE, Liu CF, Sloan KL, Malkin J, Fishman PA, Rosen AK, Loveland S, Nichol W, Suzuki NT, Perrin E, Sharp ND, Todd-Stenberg J (2003) Predicting costs of care using a pharmacy-based measure risk adjustment in a veteran population. Med Care 41(6):753–760
20. Zhao Y, Ash AS, Ellis RP, Ayanian JZ, Pope GC, Bowen B, Weyuker L (2005) Predicting pharmacy costs and other medical costs using diagnoses and drug claims. Med Care 43(1):34–43
21. Joynt KE, Orav EJ, Jha AK (2011) Thirty-day readmission rates for Medicare beneficiaries by race and site of care. JAMA 305:675–681

22. McHugh MD, Margo Brooks Carthon J, Kang XL (2010) Medicare readmissions policies and racial and ethnic health disparities: a cautionary tale. Policy Polit Nurs Pract 11(4):309–316
23. Rathore SS, Foody JM, Wang Y, Smith GL, Herrin J, Masoudi FA, Wolfe P, Havranek EP, Ordin DL, Krumholz HM (2003) Race, quality of care, and outcomes of elderly patients hospitalized with heart failure. JAMA 289:2517–2524
24. Luthi JC, Lund MJ, Sampietro-Colom L, Kleinbaum DG, Ballard DJ, McClellan WM (2003) Readmissions and the quality of care in patients hospitalized with heart failure. Int J Qual Healthcare 15(5):413–421
25. Harrell FE, Lee KL, Mark DB (1996) Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 15:361–387
26. Efron B (1983) Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc 78:316–331
27. Ross JS, Mulvey GK, Stauffer B et al (2008) Statistical models and patient predictors of readmission for heart failure: a systematic review. Arch Intern Med 168(13):1371–1386
28. Joynt KE, Jha AK (2011) Who has higher readmission rates for heart failure, and why? Implications for efforts to improve care using financial incentives. Circ Cardiovasc Qual Outcomes 4 (1):53–59
29. Williams S, Bottle A, Aylin P (2005) Length of hospital stay and subsequent emergency readmission. BMJ 331:371
30. Maddala GS (1992) Introduction to Econometrics, 2nd edn. Macmillan
31. CCS. http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. Accessed 11 January 2013
32. CMS HCC. http://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk_adjustment_prior.html. Accessed 11 January 2013
33. Anderson D, Price C, Golden B, Jank W, Wasil E (2011) Examining the discharge practices of surgeons at a large medical center. Health Care Manag Sci 14(4):338–347
34. Anderson D, Golden B, Jank W, Wasil E (2012) The impact of hospital utilization on patient readmission rate. Health Care Manag Sci 15:29–36