Published in final edited form as:
Ann Intern Med. 2014 November 18; 161(0): S66–S75. doi:10.7326/M13-3000.

Development and use of an administrative claims measure for profiling hospital-wide performance on 30-day unplanned readmission

Leora I. Horwitz, MD, MHS[1,2] **[RWJCSP alumni or current RWJCSP Program Director]**, **Chohreh Partovian, MD, PhD**[2,3], **Zhenqiu Lin, PhD**[2,3], **Jacqueline N. Grady, MS**[2], **Jeph Herrin, PhD**[3,4], **Mitchell Conover, BA**[2], **Julia Montague, MPH**[2], **Chloe Dillaway, BA**[2], **Kathleen Bartczak, BA**[2], **Lisa G. Suter, MD**[2,5] **[RWJCSP alumni or current RWJCSP Program Director]**, **Joseph S. Ross, MD, MHS**[1,2,6,7] **[RWJCSP alumni or current RWJCSP Program Director]**, **Susannah M. Bernheim, MD, MHS**[1,2] **[RWJCSP alumni or current RWJCSP Program Director]**, **Harlan M. Krumholz, MD, SM**[2,3,6,7] **[RWJCSP alumni or current RWJCSP Program Director]**, and **Elizabeth E. Drye, MD, SM**[2,8]

[1]Section of General Internal Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT

[2]Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT

[3]Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT

[4]Health Research & Educational Trust, Chicago IL

[5]Section of Rheumatology, Department of Internal Medicine, Yale School of Medicine, New Haven, CT

[6]The Robert Wood Johnson Clinical Scholars Program, Department of Internal Medicine, Yale School of Medicine, New Haven, CT

[7]Department of Health Policy and Management, Yale School of Public Health, New Haven, CT

[8]Department of Pediatrics, Yale School of Medicine, New Haven, CT

Abstract

Background—Existing publicly-reported readmission measures are condition-specific, representing < 20% of adult hospitalizations. An all-condition measure may better measure quality and promote innovation.

Corresponding author: Leora Horwitz, MD Section of General Internal Medicine P.O. Box 208093 New Haven, CT 06520-8093 Tel: (203) 688-5678.
Current mailing addresses for all authors: Leora Horwitz and Joseph Ross, Section of General Internal Medicine, P.O. Box 208093, New Haven, CT 06520
Chohreh Partovian, 109 Church Street, Unit 607, New Haven, CT 06510
Authors Lin, Grady, Montague, Suter, Ross, Bernheim, Krumholz, Drye, Center for Outcomes Research and Evaluation, 1 Church Street, Suite 200, New Haven, CT 06510
Jeph Herrin, P.O. Box 2254, Charlottesville, VA 22902
Mitchell Conover, 1227 Northcreek Drive, Durham, NC 27707
Chloe Dillaway, 10300 Jollyville Road #1031, Austin, TX 78759
Kathleen Bartczak, ul. Racławicka 9a, 30-075 Krakow, Poland

Objective—To develop an all-condition, hospital-wide readmission measure.

Design—Measure development

Setting—4,821 US hospitals.

Patients—Medicare Fee for Service (FFS) beneficiaries ≥ 65 years.

Measurements—Hospital-level, risk-standardized unplanned readmissions within 30 days of discharge. The measure uses Medicare FFS claims and is a composite of five specialty-based risk-standardized rates for medicine, surgery/gynecology, cardiorespiratory, cardiovascular and neurology cohorts. We randomly split the 2007–2008 admissions for development and validation. Models were adjusted for age, principal diagnosis and comorbidity. We examined calibration in Medicare and all-payer data, and compared hospital rankings in the development and validation samples.

Results—The development dataset contained 8,018,949 admissions associated with 1,276,165 unplanned readmissions (15.9%). The median hospital risk-standardized unplanned readmission rate was 15.8 (range 11.6–21.9). The five specialty cohort models accurately predicted readmission risk in both Medicare and all-payer datasets for average risk patients but slightly overestimated readmission risk at the extremes. Overall hospital risk-standardized readmission rates did not differ statistically in the split samples (p=0.7 for difference in rank) and 76% of hospitals' validation set rankings were within two deciles of the development rank (24% >2 deciles). Of hospitals ranking in the top or bottom deciles, 90% remained within two deciles (10% >2 deciles), and 82% remained within one decile (18% > 1 decile).

Limitations—Risk-adjustment was limited to that available in claims data.

Conclusions—We developed a claims-based hospital-wide unplanned readmission measure for profiling hospitals that produced reasonably consistent results in different datasets and was similarly calibrated in both Medicare and all-payer data.

Primary funding source—Centers for Medicare & Medicaid Services

Introduction

Readmission to the hospital within 30 days of discharge occurs for almost one-fifth of Medicare beneficiaries and costs the Centers for Medicare & Medicaid Services (CMS) $26 billion annually (1).One contributor to readmissions is the quality of the transition from hospital to home, which is often inadequate (2–5). Improvements in transitional care have been shown to reduce readmission rates (6–8). Consequently, hospital readmissions have become a key metric for efforts to promote innovations in health care that improve the quality and value of care across settings (9–11).

Rigorous development of a measure that can be used for hospital performance profiling, focused on hospital readmission rates for a broad spectrum of patients, is necessary to support these healthcare innovations. CMS publicly reports risk-standardized readmission rates (RSRRs) for heart failure, pneumonia, acute myocardial infarction and hip and knee replacement (12–15). These conditions represent less than 20% of all Medicare hospital admissions (16). While information on individual conditions is important to guide quality

improvement activities, focusing on a few conditions may not incentivize optimal distribution of resources that could be used to improve hospital-wide practices or to target different high-risk patients. Single-condition measures may limit hospitals' abilities to broadly engage physicians, staff, and community members in readmission reduction efforts. Finally, many hospitals care for few patients with each condition, necessitating multiple years of data to produce stable hospital rankings. This reduces timeliness. Thus, it is important to measure all-condition readmission rates in order to capture the majority of hospitalized patients, encourage a focus on high risk patients regardless of condition, and incentivize system- and community-wide quality improvements.

Constructing an all-condition readmission measure for profiling performance presents several challenges. The measure must account for the diversity of conditions and procedures at different hospitals, to provide a fair assessment of relative hospital performance. It must balance inclusivity (encompassing a wide range of patients) with usability (providing information that hospitals can act upon). A readmission measure should also exclude planned readmissions.

Here we describe the development of a claims-based, risk-standardized hospital-wide readmission measure which is innovative in several important respects: it includes the great majority of adult inpatients, accounts for diverse conditions and their prevalence at different institutions, excludes planned readmissions, and is a composite of 5 specialty cohort models to make the measure more informative to hospitals. This measure has been endorsed by the National Quality Forum for quality measurement, and is publicly reported by CMS (17, 18).

Methods

Data

We developed the measure under contract to CMS using hospitalizations in Medicare FFS Part A claims data. We obtained enrollment and post-discharge mortality status from the Medicare Denominator File. We developed the measure using a random half of the 2007–8 combined Medicare Provider Analysis and Review (MedPAR) data, and validated the measure using three datasets: the second half of the 2007–8 sample, the 2009 MedPAR data, and 2006 California Patient Discharge data. For each dataset, we used one prior year of inpatient data for risk adjustment.

Eligibility criteria

Qualifying index admissions must have met the following criteria: patient was admitted to a short-term acute care or critical access hospital, survived hospitalization, was age 65 years or older at discharge (18 or older when applied to general adult population), and was discharged home or to a non-short-term acute hospital setting. Multiple admissions for the same patient were included. We excluded: admissions for patients without at least 30 days of post-discharge enrollment in Medicare FFS (necessary for determining the outcome), admissions for patients not continuously enrolled in Medicare FFS during the 12 months prior to admission (necessary for risk adjustment), patients discharged against medical advice (because the hospital did not have the opportunity to provide optimal care),

admissions to Prospective Payment System (PPS)-exempt cancer hospitals (because Medicare has deemed these hospitals not comparable to other institutions), admissions for medical treatment of cancer (because of high competing mortality rates; see **Appendix A**), and admissions for rehabilitation care. We did not exclude eligible readmissions from serving as index admissions.

Rationale for the measure architecture

To optimize measure design, we explored tradeoffs between estimating risk-standardized rates for one hospital-wide cohort versus a composite measure score of rates for subgroups of patients. One model including all admissions would not be very informative for hospital improvement, and would not account for differences in the influence of risk variables across different conditions. However, hospitals did not have sufficient numbers of admissions to support separate models for all conditions. To reconcile these tensions, we tested but rejected an approach of defining 20–30 cohorts by grouping conditions with similar risk variable-outcome relationships using clustering algorithms; the resulting cohorts were clinically incoherent and sample sizes still small. Instead, we identified five cohorts organized according to service lines that were made up of conditions or procedures with relatively similar readmission and post-discharge mortality rates, that were likely to be cared for by similar teams of clinicians, and that would generate an adequate sample size for most hospitals. These cohorts included: medicine, surgery/gynecology, cardiorespiratory, cardiovascular, and neurology. Cardiorespiratory patients (e.g., heart failure, pneumonia, chronic obstructive pulmonary disease, asthma) were separated from the medicine cohort because they are clinically very similar and have the highest volume and readmission rates. This approach allowed for differential risk-adjustment, enabled sufficient sample size for most cohorts at most hospitals, and produced clinically-relevant specialty-specific results as well as a composite measure score.

To assign admissions to cohorts, we first used the Agency for Healthcare Research and Quality 2009 Clinical Classifications Software (CCS; AHRQ, Rockville, MD) (19) to group all ICD-9-based principal discharge diagnoses into one of 285 mutually exclusive condition categories, and all ICD-9-based procedure codes into one of 231 mutually exclusive procedure categories. Next we identified all procedure categories that would typically result in a patient being cared for by a surgical or gynecological service (**Appendix B**), and assigned all admissions that included one of these procedure categories to the surgery/gynecology specialty cohort. We assigned each remaining admission to one of the other four cohorts on the basis of its principal discharge diagnosis, grouped by condition category (**Appendix C**).

Outcome

The outcome was all-cause unplanned readmission to any hospital within 30 days of discharge. Because there is no code on administrative claims for identifying planned readmissions, we constructed an algorithm and refined it based on input from 27 clinical experts recommended by 15 specialty societies, and from three public comment periods (**Appendix D**). We defined planned readmissions as either: (1) readmissions for a few specific condition or procedure categories (chemotherapy/radiation therapy, organ

transplant, rehabilitation, obstetrical delivery); or (2) readmissions in which any of a list of typically-planned procedures occurred, and in which the principal diagnosis was not an acute condition or a complication of care. Readmissions not meeting either criterion were categorized as unplanned.

Risk adjustment

We adjusted both for comorbidity and for principal diagnosis. To define comorbid risk adjustment variables, we grouped ICD-9 codes into 189 CMS Condition Categories (CMS-CCs) (20) and defined a risk variable as present if it was coded in any inpatient claim in the 12 months prior to admission or as a secondary diagnosis in the index admission. For practical purposes of data processing, we elected not to include outpatient claims data. To avoid adjusting for potential complications as comorbidities, we did not code certain CMS-CCs as risk factors if they only appeared as secondary diagnosis codes in the index admission (**Appendix E**).

We began with a set of 41 variables comprised of 74 CMS-CCs based on importance in existing risk-standardized readmission models (12–14) or on clinical relevance to an all-condition measure. We ran a separate logistic regression model for each condition category, using the full set of candidate risk adjustment variables, and examined odds ratios for readmission for each variable across the different condition models. We excluded risk variables that were rarely statistically significant, and those that were not performing consistent with clinical expectations. . We then combined risk variables that were clinically coherent and carried similar risks across condition categories.

We also created indicator variables for each discharge condition category with at least 1,000 admissions yearly. All conditions with fewer than 1,000 admissions in a given specialty cohort were grouped into a single “low frequency” indicator variable. When using the California validation set, we respecified the conditions belonging to the low frequency condition groups.

Measure calculation

Using PROC GLIMMIX, we estimated a separate mixed effects logistic regression model with hospital as a random intercept for each of the five cohorts and used the results to calculate the predicted and expected numbers of readmissions at each hospital (21, 22). The predicted number of readmissions in each cohort was calculated as the sum of the predicted probability of readmission for all admissions, estimated using each hospital’s patient mix and a hospital specific effect estimated for each hospital. The hospital specific effect, also called the Empirical Bayes estimator, is an estimate of each hospital’s outcome rate; this estimate is stabilized, or “shrunk,” by pooling the adjusted rate at that hospital with the adjusted rate for all hospitals. The pooling is weighted by volume, so low volume hospitals are shifted towards the national mean more than high volume hospitals (23). The expected number of readmissions in each cohort for each hospital was similarly calculated as the sum of the predicted probability of readmission for all admissions, using each hospital’s patient mix and the average hospital-specific effect of all hospitals. We divided the predicted number of readmissions by the expected number of readmissions to obtain a standardized

readmission ratio for each specialty cohort for each hospital (21). We calculated each hospital's hospital-wide composite ratio by calculating the volume-weighted logarithmic mean of the 5 specialty cohort ratios. The logarithmic mean is the mathematically appropriate method of averaging ratios (24). Specialty cohorts with no eligible admissions at a hospital were not included in the hospital's composite. To aid in interpretation, we multiplied the composite standardized ratios by the national observed readmission rate to produce the risk-standardized hospital-wide readmission rate (RSRR). We used bootstrapping to derive an interval estimate for the final rates for each hospital (25–28) (**Appendix F**).

Statistical analysis

To assess models' discrimination, we constructed calibration curves plotting observed and predicted readmission rates for patients in each decile of predicted probability based on ordinary logistic regression models without hospital random effects, and assessed the degree of overlap (29).

Reliability is defined by the National Quality Forum as the extent to which the measure produces consistent results (30). To assess reliability of model parameters, we compared regression coefficients and standard errors of risk variables in each specialty cohort model between the development sample and each of the three validation sets. To assess the consistency of the composite hospital-wide measure, we compared hospitals' risk standardized rankings in each half of the split-sample 2007–2008 data, reporting the number of changes in deciles and the Wilcoxon signed rank statistic (31).We also plotted the differences between the two within-hospital rates against the average of the two within-hospital rates (32). Data on model c-statistics, correlation of specialty ratios with each other, and internal consistency of the composite rate appear in the technical report (18). We used SAS 9.2 (SAS Institute, Cary, NC) for analyses. The Yale University Investigational Review Board approved this study.

Role of the funding source

This work was performed under contract HHSM-500-2008-0025I/HHSM-500-T0001, Modification No. 000008, funded by CMS, an agency of the US Department of Health and Human Services. No funder had a role in analysis and interpretation of data or in writing the report. CMS advised on the study design and approved submission of the manuscript.

Results

Study cohort

The development dataset included 8,018,949 discharges from 4,821 hospitals (approximately 93% of all Medicare FFS acute care hospitalizations of patients 65 and older) (see Figure 1 for 2008 data). The mean age of the cohort was 78 years, with 58.2% women and 13.3% non-white patients. The median annual hospital volume of index admissions was 702 (interquartile range, 239, 2,246).

Specialty cohort volume ranged from 464,776 (neurology) to 3,157,943 (medicine) (Table 1). A total of 83.1% of hospitals accounting for 98.6% of admissions had at least one index admission in all five specialty cohorts.

The dataset included 1,276,165 (90.8%) unplanned and 129,436 (9.2%) planned readmissions, for an overall unplanned 30-day readmission rate of 15.9%, ranging from a minimum of 11.8% (surgery/gynecology) to a maximum of 20.7% (cardiorespiratory). In 3.9% of admissions, the patient died after discharge without being readmitted. The median risk standardized readmission rates was 15.8% (range, 11.6–21.9). Table 2 shows the distributions of rates.

Measure performance

The final 31 comorbidity variables are listed in **Appendix G**. Parameter estimates varied in magnitude but not direction across specialty cohorts (**Appendices H-L**).

Model calibration plots for the 2007–2008 Medicare split-sample development dataset, the 2009 Medicare dataset, and the 2006 California all-payer dataset are shown in Figure 2. This figure illustrates that, on a patient level, the models slightly overestimate readmission risk at the highest and lowest risks. The figure also illustrates consistent calibration of the models in 2009 Medicare data, and in all-payer data. The RSRR rank for each hospital was not significantly different between the 2007–2008 derivation and validation sets (p=0.71).

When ranked by standardized readmission rate, 76% of hospitals shifted two deciles or less between development and validation sets; put another way, 24% shifted by more than 2 deciles. Model performance was most stable at the extremes: 90% of hospitals starting in the top or bottom deciles in the derivation set shifted by two deciles or less in the validation set (10% shifted > 2 deciles), and 82% shifted one decile or less (18% > 1 decile). The difference in standardized rates between the 9^{th} and 10^{th} deciles is 1.2 percentage points, compared to a difference of 0.63 percentage points between the central 4^{th} and 7^{th} deciles.

Figure 3 cross classifies the within hospital differences in risk standardized rates between the derivation and validation sets against the within-hospital means. Ninety-five percent of hospitals have a difference of less than 1.4 percentage points, and outlier differences are nearly all among hospitals of average performance, indicated by the two vertical 95% confidence interval lines. Hospitals falling in or near areas I, III, VII, and IX are those with extreme rates that varied more than average between the datasets.

Discussion

We developed a hospital-wide, 30-day unplanned readmission measure that is risk-standardized to account for differences in comorbidity and in the distribution of diagnoses within each hospital, and can be used to measure hospital performance. The measure had reasonable split-sample consistency in Medicare data, and performed well in subsequent years of data as well as in the full adult (18 years and older) patient population. It broadens the scope of readmission outcome measurement from a minority of primarily medical patients to the vast majority of a hospital's patients, including those cared for by surgeons,

neurologists, gynecologists and others, thus providing a more comprehensive view of readmissions. Unlike other all-condition readmission measures, it excludes planned readmissions and is comprised of multiple clinically-distinct rates to increase usability. Finally, the measure conforms to standards for publicly-reported outcome measures (34). CMS began publicly reporting this measure in December, 2013.

The simultaneous all-condition and specialty-specific nature of this measure makes it particularly suitable for helping institutions identify areas needing improvement. The hospital-wide rates apply to over 90% of admissions, making the measure broadly applicable. This global rate can be publicly reported and benchmarked against national averages, enabling patients, payers and clinicians to select hospitals based on results, and incentivizing poorly-performing hospitals to improve (35–37). In addition, the measure produces specialty-specific rates that can be provided to hospitals confidentially to help them identify care teams or patient populations for particular focus. In this way the measure takes a unique approach of providing both an overall incentive for change and more specific data to target change efforts.

An additional novel feature of this measure is the planned readmission algorithm, vetted through extensive expert consultations and public comment (38), which enabled us to exclude planned readmissions. Some measures have attempted to count only readmissions that are “preventable” or “related” to the index admission, on the assumption that any “unrelated” readmission is necessarily unpreventable (39–41). However, “unrelated” readmissions may be consequences of stressors during hospitalization (42) or low quality of care provided during the index admission (3); conversely, some “related” readmissions are unavoidable due to natural progression of disease. Furthermore, there is little evidence to suggest that “relatedness” can reliably be determined even with detailed chart review (43, 44). Instead, we counted all readmissions except those that were likely to have been planned. In doing so we acknowledge that the ideal readmission rate is not zero: many patients will unavoidably be readmitted. The measure assumes that the proportion of unavoidable readmissions should be similar across hospitals given similar care quality, once we account for case mix and principal diagnoses. Excluding planned readmissions from the measure creates an opportunity for gaming; however, as the planned readmissions are largely identified through procedures that are performed during the readmission, we anticipate that opportunity for gaming will be limited. CMS conducts routine surveillance for evidence of unintended consequences and if necessary, measure specifications can be altered in response.

We have identified eight other all-condition readmission measures, three of which have been used in the US or Canada, though none currently on a national scale (39–41, 45–51). All but two use a similar 28 or 30-day timeframe. Some include virtually all patients (46) and others include only a narrow spectrum (41), but all exclude transfers to acute settings and hospitalizations in which the patient died, as does this measure. Three of the measures, like ours, also exclude patients admitted for cancer treatment and those who left against medical advice (40, 41, 48). Some measures exclude planned readmissions (45, 46, 48). Nearly all the other measures use risk adjustment, but none uses models appropriate for clustered data as recommended by outcome measure guidelines.(34, 52)

This measure was designed to profile hospital quality by benchmarking hospitals against national performance. As such it may catalyze improvement activities, and can be used to track national trends over time, but it cannot be replicated by individual hospitals, which lack access to national data. It was not designed to track internal improvements (for which risk-standardization against national data is not necessary), nor as a tool to predict individual patient readmission risk. We deliberately do not include covariates such as race, income, previous admission, complications during hospitalization or length of stay even though they may improve patient-level prediction (53) because they may represent variation in the outcome due to hospital practice that the measure is intended to capture. Adjusting for race or income might obscure differences in care provided to these patients. We do not want to adjust for poor quality when trying to measure quality. Notably, only one of the eight other all-condition readmission measures includes race as a covariate (48) and none includes income, education, previous admission or length of stay. For example, adjusting for complications would perversely give credit to hospitals with more readmissions caused by complications of care. We do not adjust for previous admission because repeated admissions may be an indicator of failed transitions, inadequate attention to goals of care, or other gaps in hospital- and community-level care.

When benchmarking hospitals, it is important to ensure results are reproducible and not unduly subject to random differences in patient mix, variation in measurement or patient coding, or unexplained random variation. We found no statistical difference in rank ordering of hospitals between randomly split development and validation sets and three quarters of hospitals moved fewer than 3 deciles between the development and validation datasets. Nonetheless, 24% of hospitals moved 3 deciles or more. In this regard, it is important to consider the limitations of a rank order analysis. First, moves among middle deciles are small and may not be as clinically meaningful as moves in extreme deciles. However, moves among middle deciles occurred more frequently than among outlier deciles. We found that outlier ranks were more consistent, with 90% of hospitals in the top or bottom deciles remaining in the top or bottom 3 deciles in the validation set. These findings should be considered in using the measure score in profiling. In addition, the number of hospitals changing ranks depends on the number of ranks selected; we chose a conservative decile approach. Further, since deciles are divided based on single points on a continuous scale, hospitals close to the dividing lines will naturally move decile ranks even with virtually identical results. Most importantly, a simple rank order ignores sampling error because it does not incorporate confidence intervals. For these reasons, the measure is publicly reported based on statistical outlier status, not ranks. The stability of outlier status has not yet been established and will be an important focus for future work in this area.

Our measure has several limitations. First, there is no gold standard against which to compare this measure to assess validity. Second, it is based on administrative data, which are known to contain errors, and which do not include specific information on disease severity. However, claims data are more complete and easily obtained than chart data, and prior readmission measures based on claims data were shown to have good agreement with measures based on chart data (25–27). Third, it has only fair patient-level predictive capacity, though our assessment of patient-level discrimination may be hampered by clustering within hospitals. Fourth, using a one year look-back for comorbidities may

artificially make patients at high intensity hospitals appear sicker;(54) on the other hand, using only comorbidities identified during the index admission would undercount risk for truly sicker patients. Fifth, planned readmissions are not explicitly flagged in administrative data, requiring us to develop an algorithm to identify them. Although it is not possible to perfectly identify planned readmissions using claims data, we used input from many surgical experts and the public to improve the algorithm and adequately trade off precision vs. usability. Sixth, competing mortality is always a concern in readmission measures, although we minimized this risk by excluding conditions with the highest competing mortality. Finally, readmission risk is also influenced by community factors such as access to care, local practice patterns and sociodemographics. Therefore this measure should be considered in conjunction with complementary measures such as community-level admission and readmission rates, and mortality. Nonetheless, it will remain critically important to measure hospital performance both to identify problems and to catalyze hospital-community partnerships.

This measure reports risk-standardized readmission rates for over 90% of admissions to acute care hospitals. It performs well in both Medicare and all-payer data and has reasonably stable performance over time. The structure of the measure, which includes separate models for specialty cohorts, increases its usability for hospital quality improvement while still producing summary results for consumers. Ultimately, the utility of the measure will depend on the degree to which hospitals and communities can work together to reduce unnecessary hospital readmission.

Acknowledgments

Financial support: This work was performed under contract HHSM-500-2008-0025I/HHSM-500-T0001, Modification No. 000008, entitled “Measure Instrument Development and Support,” funded by CMS, an agency of the US Department of Health and Human Services. Dr. Horwitz is supported by the National Institute on Aging (K08 AG038336) and by the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program. Dr. Horwitz is also a Pepper Scholar with support from the Claude D. Pepper Older Americans Independence Center at Yale University School of Medicine (#P30AG021342 NIH/NIA). Dr. Ross is supported by the National Institute on Aging (K08 AG032886) and by the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program. Dr. Krumholz is supported by grant U01 HL105270-03 (Center for Cardiovascular Outcomes Research at Yale University) from the National Heart, Lung, and Blood Institute. No funder had a role in analysis and interpretation of data or in writing the report. CMS advised on the study design and approved submission of the manuscript.

We gratefully acknowledge the support and guidance of Lein Han, PhD and Michael Rapp, MD, statistical guidance from Sharon-Lise Normand, PhD and research assistance from Chinwe Nwosu. This work was performed under contract HHSM-500-2008-0025I/HHSM-500-T0001, Modification No. 000008, entitled “Measure Instrument Development and Support,” funded by CMS, an agency of the US Department of Health and Human Services. Dr. Horwitz is supported by the National Institute on Aging (K08 AG038336) and by the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program. Dr. Horwitz is also a Pepper Scholar with support from the Claude D. Pepper Older Americans Independence Center at Yale University School of Medicine (#P30AG021342 NIH/NIA). Dr. Ross is supported by the National Institute on Aging (K08 AG032886) and by the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program. Dr. Krumholz is supported by grant U01 HL105270-03 (Center for Cardiovascular Outcomes Research at Yale University) from the National Heart, Lung, and Blood Institute. No funder had a role in analysis and interpretation of data or in writing the report. CMS advised on the study design and approved submission of the manuscript.The content is solely the responsibility of the authors and does not necessarily represent the official views or policies of the US Department of Health and Human Services, the National Institute on Aging, the National Heart, Lung, and Blood Institute, or the American Federation for Aging Research.

References

1. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. N Engl J Med. 2009; 360(14):1418–28. [PubMed: 19339721]
2. Bradley EH, Curry L, Horwitz LI, Sipsma H, Thompson JW, Elma M, et al. Contemporary evidence about hospital strategies for reducing 30-day readmissions: a national study. J Am Coll Cardiol. 2012; 60(7):607–14. [PubMed: 22818070]
3. Ziaeian B, Araujo KL, Van Ness PH, Horwitz LI. Medication reconciliation accuracy and patient understanding of intended medication changes on hospital discharge. J Gen Intern Med. 2012; 27(11):1513–20. [PubMed: 22798200]
4. Horwitz LI, Jenq GY, Brewster UC, Chen C, Kanade S, Van Ness PH, et al. Comprehensive quality of discharge summaries at an academic medical center. J Hosp Med. 2013; 8(8):436–43. [PubMed: 23526813]
5. Horwitz LI, Moriarty JP, Chen C, Fogerty RL, Brewster UC, Kanade S, et al. Quality of discharge practices and patient understanding at an academic medical center. JAMA Intern Med. 2013; 173(18):1715–22. [PubMed: 23958851]
6. Naylor MD, Brooten D, Campbell R, Jacobsen BS, Mezey MD, Pauly MV, et al. Comprehensive discharge planning and home follow-up of hospitalized elders: a randomized clinical trial. Jama. 1999; 281(7):613–20. [PubMed: 10029122]
7. Coleman EA, Parry C, Chalmers S, Min SJ. The care transitions intervention: results of a randomized controlled trial. Arch Intern Med. 2006; 166(17):1822–8. [PubMed: 17000937]
8. Jack BW, Chetty VK, Anthony D, Greenwald JL, Sanchez GM, Johnson AE, et al. A reengineered hospital discharge program to decrease rehospitalization: a randomized trial. Ann Intern Med. 2009; 150(3):178–87. [PubMed: 19189907]
9. Medicare Shared Savings Program, Patient Protection and Affordable Care Act, Pub. L. No. 124, §3022 (March 23, 2010).
10. Hospital Readmission Reduction Program, Patient Protection and Affordable Care Act, Pub. L. No. 124, §3025 (March 23, 2010).
11. Community Care Transitions Program, Patient Protection and Affordable Care Act, Pub. L. No. 124, §3026 (March 23, 2010).
12. Keenan PS, Normand SL, Lin Z, Drye EE, Bhat KR, Ross JS, et al. An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. Circ Cardiovasc Qual Outcomes. 2008; 1:29–37. [PubMed: 20031785]
13. Krumholz HM, Lin Z, Drye EE, Desai MM, Han LF, Rapp MT, et al. An administrative claims measure suitable for profiling hospital performance based on 30-day all-cause readmission rates among patients with acute myocardial infarction. Circ Cardiovasc Qual Outcomes. 2011; 4(2): 243–52. [PubMed: 21406673]
14. Lindenauer PK, Normand SL, Drye EE, Lin Z, Goodrich K, Desai MM, et al. Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia. J Hosp Med. 2011; 6(3):142–50. [PubMed: 21387551]
15. Grosso, LM.; Curtis, JP.; Z, L.; Geary, LL.; Vellanky, S.; Oladele, C., et al. [25 March 2014] Hospital-level 30-day all-cause risk-standardized readmission rate following elective primary total hip arthroplasty (THA) and/or total knee arthroplasty (TKA). Accessed at http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1219069855841 on
16. Centers for Disease Control and Prevention. [29 August 2013] National Hospital Discharge Survey. Accessed at http://www.cdc.gov/nchs/nhds.htm on
17. Changes to the Hospital Inpatient and Long-Term Care Prospective Payment System for FY 2013 (CMS-1588-P). Federal Register. 2012; 77(170):53521.
18. Horwitz, L.; Partovian, C.; Lin, Z.; Herrin, J.; Grady, J.; Conover, M., et al. [28 August 2013] Hospital-wide all-cause unplanned readmission measure: Final technical report. Accessed at https://www.qualitynet.org/dcs/ContentServer?cid=1219069855273&pagename=QnetPublic%2FPage%2FQnetTier3&c=Page on

19. Healthcare Cost and Utilization Project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality; 2006–2009. HCUP Clinical Classifications Software (CCS) for ICD-9-CM.
20. Pope GC, Kautter J, Ellis RP, Ash AS, Ayanian JZ, Lezzoni LI, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financ Rev. 2004; 25(4):119–41. [PubMed: 15493448]
21. Normand SLT, Glickman ME, Gatsonis CA. Statistical methods for profiling providers of medical care: Issues and applications. J Am Stat Assoc. 1997; 92(439):803–14.
22. He Y, Selck F, Normand SL. On the accuracy of classifying hospitals on their performance measures. Statistics in medicine. 2014; 33(7):1081–103. [PubMed: 24122879]
23. Ash, AS.; Fienberg, SE.; Louis, TA.; Normand, SL.; Stukel, TA.; Utts, J. Statistical Issues in Assessing Hospital Performance: Commissioned by the Committee of Presidents of Statistical Societies. 2012.
24. Fleming PJ, Wallace JJ. How Not to Lie with Statistics - the Correct Way to Summarize Benchmark Results. Communications of the Acm. 1986; 29(3):218–21.
25. Krumholz, H.; Normand, SL.; Keenan, P.; Desai, M.; Lin, Z.; Drye, E., et al. [21 March 2014] Hospital 30-Day AMI Readmission Measure Methodology: Report prepared for the Centers for Medicare & Medicaid Services. Accessed at http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1219069855841 on
26. Krumholz, H.; Normand, SL.; Keenan, P.; Lin, Z.; Drye, E.; Bhat, K., et al. [21 March 2014] Hospital 30-Day Heart Failure Readmission Measure Methodology: Report prepared for the Centers for Medicare & Medicaid Services. Accessed at http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1219069855841 on
27. Krumholz, H.; Normand, SL.; Keenan, P.; Desai, M.; Lin, Z.; Drye, E., et al. [21 March 2014] Hospital 30-Day Pneumonia Readmission Measure Methodology: Report prepared for the Centers for Medicare & Medicaid Services. Accessed at http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1219069855841 on
28. Normand SL, Wang Y, Krumholz HM. Assessing surrogacy of data sources for institutional comparisons. Health Serv Outcomes Res Method. 2007; 7(1-2):79–96.
29. Harrell, FE. Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; 2002.
30. National Quality Forum. [24 March 2014] Measure Evaluation Criteria. Accessed at https://www.qualityforum.org/docs/measure_evaluation_criteria.aspx#scientific on
31. Wilcoxon F. Individual comparisons by ranking methods. Biometrics Bulletin. 1945; 1(6):80–3.
32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1(8476):307–10. [PubMed: 2868172]
33. Landis JKG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:15.
34. Krumholz HM, Brindis RG, Brush JE, Cohen DJ, Epstein AJ, Furie K, et al. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. Circulation. 2006; 113(3):456–62. [PubMed: 16365198]
35. Mukamel DB, Mushlin AI. Quality of care information makes a difference: an analysis of market share and price changes after publication of the New York State Cardiac Surgery Mortality Reports. Med Care. 1998; 36(7):945–54. [PubMed: 9674613]
36. Kaiser Health News. [November 11, 2013] Latest Destination For Medical Tourism: The US. Accessed at http://www.kaiserhealthnews.org/stories/2010/july/07/domestic-medical-tourism.aspx on
37. Kaiser Health News. [November 11, 2013] Wal-Mart Expanding Program For No-Cost Employee Surgeries. Accessed at http://www.kaiserhealthnews.org/Daily-Reports/2012/October/12/marketplace-walmart.aspx on
38. Centers for Medicare and Medicaid Services. [21 March 2014] Hospital Wide Readmissions Verbatim Comments Report. Accessed at http://www.cms.gov/Medicare/Quality-Initiatives-

Patient-Assessment-Instruments/MMS/downloads/MMSHospitalWideReadmissionVerbatimCommentsReport.pdf on

39. Halfon P, Eggli Y, Pretre-Rohrbach I, Meylan D, Marazzi A, Burnand B. Validation of the potentially avoidable hospital readmission rate as a routine indicator of the quality of hospital care. Med Care. 2006; 44(11):972–81. [PubMed: 17063128]
40. Goldfield NI, McCullough EC, Hughes JS, Tang AM, Eastman B, Rawlins LK, et al. Identifying potentially preventable readmissions. Health Care Financing Review. 2008; 30(1):75–91. [PubMed: 19040175]
41. Anderson, GM.; Brown, AD.; Doran, D.; Howe, N.; Green, J.; Tallentire, M. [21 March 2014] Hospital e-Scorecard Report 2008: Acute Care Clinical Utilization and Outcomes Technical Summary. Accessed at https://ozone.scholarsportal.info/bitstream/1873/13453/1/288309.pdf on
42. Krumholz HM. Post-hospital syndrome--an acquired, transient condition of generalized risk. N Engl J Med. 2013; 368(2):100–2. [PubMed: 23301730]
43. van Walraven C, Jennings A, Taljaard M, Dhalla I, English S, Mulpuru S, et al. Incidence of potentially avoidable urgent readmissions and their relation to all-cause urgent readmissions. CMAJ. 2011; 183(14):E1067–72. [PubMed: 21859870]
44. van Walraven C, Jennings A, Forster AJ. A meta-analysis of hospital 30-day avoidable readmission rates. J Eval Clin Pract. 2011; 18(6):1211–8. [PubMed: 22070191]
45. Chambers M, Clarke A. Measuring readmission rates. BMJ. 1990; 301(6761):1134–6. [PubMed: 2132558]
46. DesHarnais S, McMahon LF Jr, Wroblewski R. Measuring outcomes of hospital care using multiple risk-adjusted indexes. Health Services Research. 1991; 26(4):425–45. [PubMed: 1917500]
47. DesHarnais S, Hogan AJ, McMahon LF Jr, Fleming S. Changes in rates of unscheduled hospital readmissions and changes in efficiency following the introduction of the Medicare prospective payment system. An analysis using risk-adjusted data. Eval Health Prof. 1991; 14(2):228–52. [PubMed: 10111358]
48. Wray NP, Peterson NJ, Souchek J, Ashton CM, Hollingsworth JC. Application of an analytic model to early readmission rates within the Department of Veterans Affairs. Medical Care. 1997; 35(8):768–81. [PubMed: 9268250]
49. UnitedHealthcare. [29 August 2013] Risk-Adjusted 30-Day All-Cause Readmission Rate. Accessed at http://www.uhc.com/physicians/practice_resources/nqf_readmission_measure_resubmission.htm on
50. The National Committee for Quality Assurance. [29 August 2013] Plan All-Cause Readmissions. Accessed at http://www.qualityforum.org/QPS/MeasureDetails.aspx?standardID=1768&print=1&entityTypeID=1 on
51. Canadian Broadcasting Corporation. [3 Sep 2013] Rate my hospital. Accessed at http://www.cbc.ca/news/health/features/ratemyhospital on
52. National Quality Forum. National voluntary consensus standards for patient outcomes, first report for phases 1 and 2: A consensus report. 2010.
53. Amarasingham R, Moore BJ, Tabak YP, Drazner MH, Clark CA, Zhang S, et al. An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data. Med Care. 2010; 48(11):981–8. [PubMed: 20940649]
54. Wennberg JE, Staiger DO, Sharp SM, Gottlieb DJ, Bevan G, McPherson K, et al. Observational intensity bias associated with illness adjustment: cross sectional analysis of insurance claims. BMJ. 2013:346.

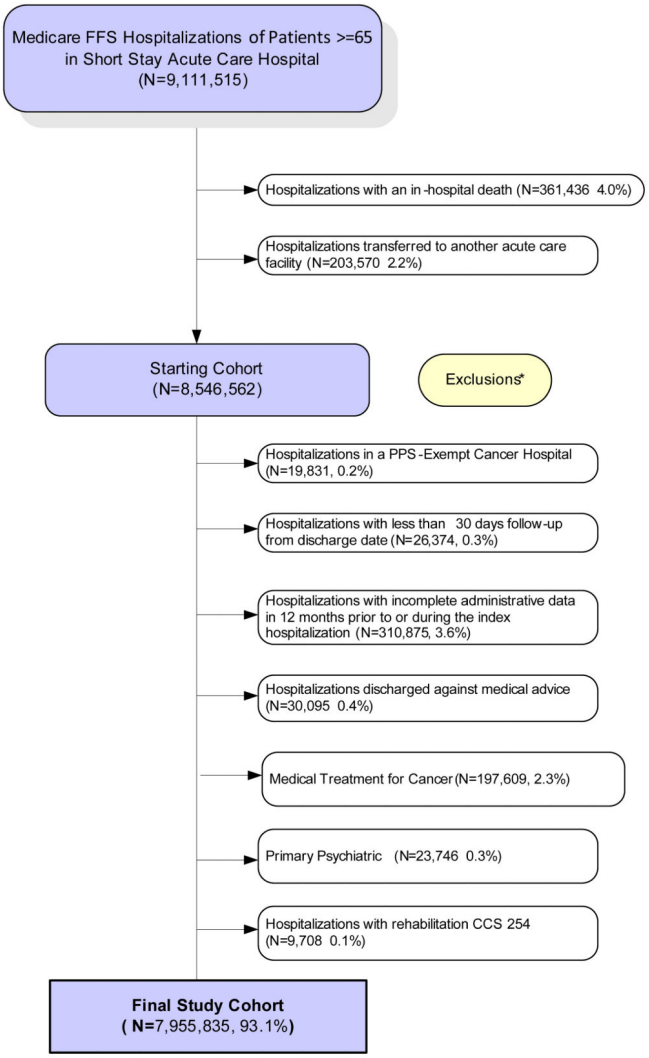

Figure 1.

Flow diagram of inclusion and exclusion criteria applied to 2008 MedPAR data.

FFS: Fee for service; PPS: prospective payment service; CCS: clinical classification software

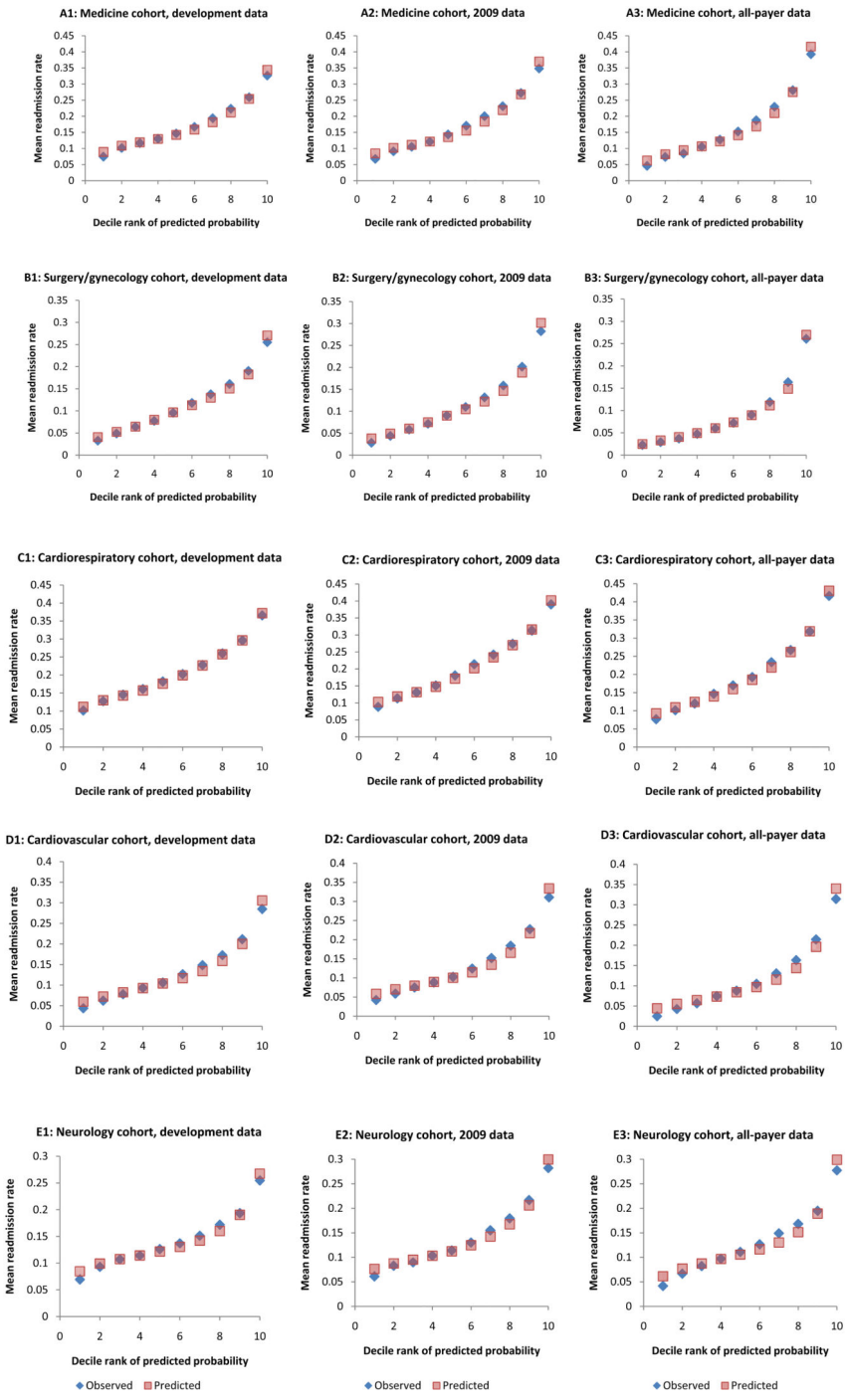

Figure 2.

Observed and predicted readmission rates for patients in each decile of predicted probability Calibration plots, by cohort, for development data set (2007–2008 split sample), Medicare 2009 data, and California 2006 all-payer data.

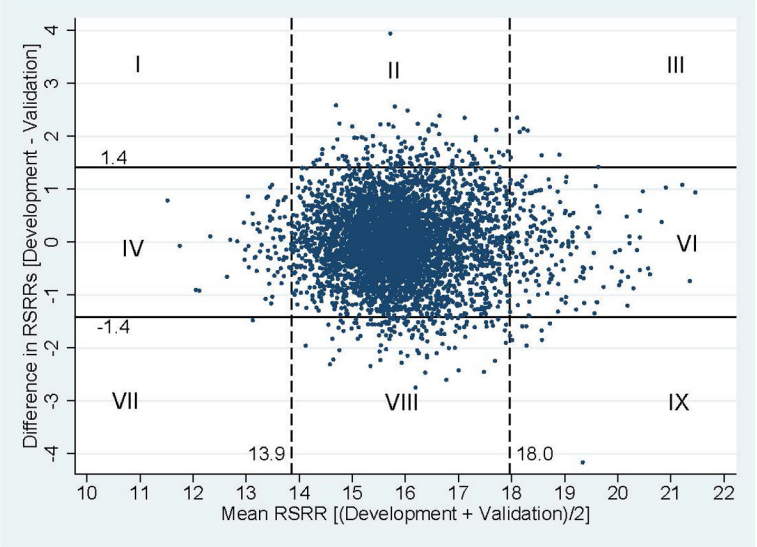

Figure 3.

Agreement of development and validation risk standardized readmission rates.

Plot of the difference between the development and validation risk standardized readmission rates (RSRRs) against the average of the two. Horizontal and vertical lines reflect the bounds of 95% of the hospitals. The center box is area V. Hospitals in or near areas I, III, VII, and IX reflect those institutions with extreme rates that tend to vary substantially between the two datasets.

Table 1

Admissions, readmissions and mortality for the five cohorts (2007–2008 development dataset)

Specialty cohort	Admissions (A)	30-day unplanned readmissions (R)	Unadjusted 30- day unplanned rate (=R/A)	30-day post-discharge mortality without readmission (M)	30-day post-discharge mortality rate readmissions without readmission (=M/A)	Planned readmission (P)	Unadjusted planned 30-day readmission rate (=P/A)	Percent of all readmissions that are planned (=P/R)
Medicine	3,157,943	549,345	17.4%	154,855	4.9%	51,408	1.6%	8.6%
Surgery/gynecology	1,889,282	223,071	11.8%	32,875	1.7%	22,269	1.2%	9.1%
Cardiorespiratory	1,413,209	292,606	20.7%	74,753	5.3%	14,397	1.0%	4.7%
Cardiovascular	1,093,739	145,201	13.3%	23,568	2.2%	35,367	3.2%	19.6%
Neurology	464,776	65,942	14.2%	29,986	6.5%	5,995	1.3%	8.3%
Total	**8,018,949**	**1,276,165**	**15.9%**	**316,037**	**3.9%**	**129,436**	**1.6%**	**9.2%**

Table 2

Hospital-level unadjusted and risk standardized readmission rates (development dataset)

Variable	N	Unadjusted mean (SD)	Unadjusted median (IQR)	Mean RSRR (SD)	Median RSRR (IQR)
Medicine	4,942	16.4 (6.5)	16.4 (13.9, 18.8)	17.5 (1.6)	17.3 (16.5, 18.3)
Surgery/gynecology	4,343	11.7 (9.3)	11.1 (8.1, 14.1)	11.8 (1.0)	11.8 (11.3, 12.3)
Cardiovascular	4,711	14.2 (8.4)	13.5 (10.5, 17.1)	13.3 (0.8)	13.3 (12.9, 13.7)
Cardiorespiratory	4,808	19.4 (6.9)	19.5 (16.1, 22.7)	20.8 (1.7)	20.6 (19.7, 21.7)
Neurology	4,691	13.7 (10.3)	13.1 (9.1, 17.1)	14.2 (1.0)	14.1 (13.7, 14.6)
HWR	4,997	15.6 (5.6)	15.5 (13.2, 18.0)	15.9 (1.1)	15.8 (15.2, 16.4)

SD: standard deviation; IQR: interquartile range; SRR: standardized readmission ratio; HWR: hospital-wide readmission rate