

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b6)

Measure Title: [Pediatric Lower Respiratory Infection Readmission Measure](#)

Date of Submission: [2/5/2014](#)

Type of Measure:

| | |
|---|--|
| <input type="checkbox"/> Composite – STOP – use composite testing form | <input checked="" type="checkbox"/> Outcome (including PRO-PM) |
| <input type="checkbox"/> Cost/resource | <input type="checkbox"/> Process |
| <input type="checkbox"/> Efficiency | <input type="checkbox"/> Structure |

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- For outcome and resource use measures, section 2b4 also must be completed.**
- If specified for **multiple data sources/sets of specifications** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 20 pages (*including questions/instructions*; minimum font size 11 pt; do not change margins). **Contact NQF staff if more pages are needed.**
- Contact NQF staff regarding questions. Check for resources at [Submitting Standards webpage](#).

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing ¹⁰ demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise.

2b2. Validity testing ¹¹ demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; ¹²

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). ¹³

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors that influence the measured outcome (but not factors related to disparities in care or the quality of care)

and are present at start of care; ^{14,15} and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** ¹⁶ **differences in performance;**

OR

there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.

15. Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care, such as race, socioeconomic status, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer or inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out the differences.

16. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of \$25 in cost for an episode of care (e.g., \$5,000 v. \$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing, (e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation. **If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.**)

| Measure Specified to Use Data From: (must be consistent with data sources entered in S.23) | Measure Tested with Data From: |
|---|---|
| <input type="checkbox"/> abstracted from paper record | <input type="checkbox"/> abstracted from paper record |
| <input checked="" type="checkbox"/> administrative claims | <input checked="" type="checkbox"/> administrative claims |
| <input type="checkbox"/> clinical database/registry | <input type="checkbox"/> clinical database/registry |
| <input type="checkbox"/> abstracted from electronic health record | <input type="checkbox"/> abstracted from electronic health record |
| <input type="checkbox"/> eMeasure (HQMF) implemented in EHRs | <input type="checkbox"/> eMeasure (HQMF) implemented in EHRs |
| <input type="checkbox"/> other: Click here to describe | <input type="checkbox"/> other: Click here to describe |

1.2. If an existing dataset was used, identify the specific dataset (the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry).

We developed and tested the measure using Medicaid Analytic eXtract (MAX) data for 26 states, which include Medicaid claims from children's and non-children's hospitals.

1.3. What are the dates of the data used in testing?

We used MAX data for hospitalizations with discharge dates from December 1, 2007 to December 31, 2008.

1.4. What levels of analysis were tested? (testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan)

| Measure Specified to Measure Performance of: (must be consistent with levels entered in item S.26) | Measure Tested at Level of: |
|---|--|
| <input type="checkbox"/> individual clinician | <input type="checkbox"/> individual clinician |
| <input type="checkbox"/> group/practice | <input type="checkbox"/> group/practice |
| <input checked="" type="checkbox"/> hospital/facility/agency | <input checked="" type="checkbox"/> hospital/facility/agency |
| <input type="checkbox"/> health plan | <input type="checkbox"/> health plan |
| <input type="checkbox"/> other: | <input type="checkbox"/> other: |

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample)

The MAX dataset includes 1,743 hospitals with ≥ 1 pediatric hospitalization for LRI. The median hospital volume of annual pediatric LRI index hospitalizations for these hospitals was 58 (IQR 18-189). Characteristics of these hospitals are detailed in the table below.

Table 1 – MAX Cohort Hospital Characteristics (Total N = 1,743)

| Hospital Characteristics | Hospitals | Index Hospitalizations |
|--|-------------|------------------------|
| Children's hospital [Number (%) of hospitals] | 78 (4.5%) | 14,734 (21.9%) |
| Teaching hospital [Number (%) of hospitals] | 114 (6.6%) | 15,769 (23.5%) |
| Rural/urban location [Number (%) of hospitals] | | |
| Urban | 608 (34.9%) | 44,561 (66.4%) |
| Suburban | 88 (5.1%) | 992 (1.5%) |
| Large town | 390 (22.4%) | 13,198 (19.7%) |
| Small town | 455 (26.1%) | 6,522 (9.7%) |
| Rural | 200 (11.5%) | 1,865 (2.8%) |

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample)

There were 1,738,043 records for pediatric patients (those ≤ 18 years, 29 days old on admission) in the MAX dataset with discharge dates from December 1, 2007 to December 31, 2008. We provide below the number and percentage of records for pediatric patients excluded because they met the indicated exclusion criteria for the measure (i.e., these records were excluded from both index hospitalizations and readmissions):

1. The hospital was a specialty or non-acute care hospital: 212,593 (12.2%)
2. Records for the hospitalization contain incomplete data for variables needed to assess eligibility for the measure or calculate readmission rates, including hospital type, patient identifier, admission date, discharge date, disposition status, date of birth, primary diagnosis code, or gender: 174,724 (10.0%)
3. The hospitalization was for birth of a healthy newborn: 599,419 (34.5%)
4. The hospitalization was for obstetric care, including labor and delivery: 56,281 (3.2%)
5. The primary diagnosis code was for a mental health condition: 25,459 (1.5%)
6. Information for some hospitalizations is contained in multiple records. These records were combined into a single record for each hospitalization, reducing the total number of records: 142,720 (8.2%)

We indicate below the number and percentage of records excluded from *index hospitalizations only* because they met the indicated exclusion criteria for index hospitalizations:

1. The patient was 18 years old or greater at the time of discharge: 984 (0.1%)
2. The discharge disposition was death: 11 (0.001%)
3. The discharge disposition was an outcome other than discharged or death (e.g., left against medical advice): 7,976 (0.5%)
4. The hospital had incomplete data or was located in a state not being analyzed: 15,038 (0.9%)
5. Thirty days of follow-up data are not available for assessing readmissions because the discharge date of the hospitalization occurred in the last month of the dataset: 18,614 (1.1%)
6. Thirty days of follow-up data are not available for assessing readmissions because the patient was enrolled in Medicaid for < 30 days after discharge from the index hospitalization: 60,220 (3.5%)
7. A hospitalization that occurs within 30 days of an index hospitalization was not counted as a new index hospitalization: 2,128 (0.1%)

8. The hospitalization does not have a primary LRI diagnosis or does not have a secondary LRI diagnosis plus a primary diagnosis of asthma, respiratory failure, or sepsis/bacteremia: 354,685 (20.4%)

After applying all of the above exclusions, 67,191 index hospitalizations remained for patients whose characteristics are described in the table below.

Table 2 – MAX Cohort Patient Characteristics (Total N = 67,191)

| Patient Characteristic | | Number (%) of Index Hospitalizations |
|-------------------------------------|---|--------------------------------------|
| Age | < 1 year | 33,175 (49.4) |
| | 1 to < 5 years | 24,715 (36.8) |
| | 5 to < 8 years | 4,501 (6.7) |
| | 8 to <12 years | 2,544 (3.8) |
| | 12 to < 18 years | 2,257 (3.4) |
| Gender | Female | 29,008 (43.2) |
| Chronic Condition Indicators (CCIs) | CCI 1 - Infectious and parasitic disease | 27 (0.04) |
| | CCI 2 - Neoplasms | 224 (0.3) |
| | CCI 3 - Endocrine, nutritional, and metabolic diseases and immunity disorders | 1,762 (2.6) |
| | CCI 4 - Diseases of blood and blood-forming organs | 2,464 (3.7) |
| | CCI 5 - Mental disorders | 1,620 (2.4) |
| | CCI 6 - Diseases of the nervous system and sense organs | 2,586 (3.8) |
| | CCI 7 - Diseases of the circulatory system | 1,115 (1.7) |
| | CCI 8 - Diseases of the respiratory system | 18,377 (27.4) |
| | CCI 9 - Diseases of the digestive system | 2,756 (4.1) |
| | CCI 10 - Diseases of the genitourinary system | 123 (0.2) |
| | CCI 12 - Diseases of the skin and subcutaneous tissue | 213 (0.3) |
| | CCI 13 - Diseases of the musculoskeletal system | 246 (0.4) |
| | CCI 14 - Congenital anomalies | 3,798 (5.6) |
| | CCI 15 - Certain conditions originating in the perinatal period | 27 (0.04) |
| | CCI 16 - Symptoms, signs, and ill-defined conditions | 114 (0.2) |
| | CCI 17 - Injury and poisoning | 6 (0.01) |
| | CCI 18 - Factors influencing health status and contact with health services | 1,924 (2.9) |
| CCI count | 0 or 1 body system | 60,576 (90.1) |
| | 2 body systems | 4,204 (6.3) |
| | 3 body systems | 1,567 (2.3) |
| | 4+ body systems | 844 (1.3) |

| | | |
|-----------------------|------------------------|---------------|
| Race/ethnicity | Asian/Pacific Islander | 861 (1.4) |
| | Black | 14,051 (22.6) |
| | Latino | 19,444 (31.2) |
| | Mixed | 520 (0.8) |
| | Native American | 3,403 (5.5) |
| | White | 24,010 (38.6) |
| Rural/urban residence | Urban | 37,235 (55.6) |
| | Suburban | 4,839 (7.2) |
| | Large town | 11,867 (17.7) |
| | Small town | 7,899 (11.8) |
| | Rural | 5,109 (7.6) |

The distribution by state of the LRI index hospitalizations is shown in the table below.

Table 3 – LRI Index Hospitalizations by State (Total N=67,191)

| State | Index Admissions | Percentage |
|----------------|------------------|------------|
| Alabama | 1,337 | 2.0% |
| Arizona | 4,837 | 7.2% |
| Connecticut | 119 | 0.2% |
| Iowa | 1,283 | 1.9% |
| Idaho | 719 | 1.1% |
| Indiana | 1,931 | 2.9% |
| Kansas | 1,662 | 2.5% |
| Kentucky | 2,184 | 3.2% |
| Louisiana | 4,776 | 7.1% |
| Minnesota | 1,332 | 2.0% |
| Missouri | 3,207 | 4.8% |
| Mississippi | 4,201 | 6.2% |
| Montana | 404 | 0.6% |
| North Carolina | 4,465 | 6.6% |
| North Dakota | 248 | 0.4% |
| New Jersey | 2,283 | 3.4% |
| New Mexico | 1,486 | 2.2% |
| New York | 7,369 | 11.0% |
| Oklahoma | 3,733 | 5.6% |
| Oregon | 841 | 1.2% |
| South Dakota | 642 | 1.0% |
| Texas | 14,376 | 21.4% |
| Virginia | 1,643 | 2.4% |
| Vermont | 126 | 0.2% |
| Wisconsin | 1,533 | 2.3% |
| Wyoming | 454 | 0.7% |

In addition, we indicate below the number and percentage of hospitalizations excluded *from readmissions only* because they met the indicated exclusion criteria for readmissions:

1. Hospitalizations with a primary ICD-9-CM procedure code for a planned procedure: 190 (4.7%) of 4,019 readmissions

2. Hospitalizations with a primary chemotherapy v-code or a primary chemotherapy procedure code: 15 (0.4%) of 4,019 readmissions

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

To evaluate the criterion validity of the measure—its ability to identify correctly the outcome of interest, readmission—we assessed performance of the measure against the gold standard of chart reviews. To perform this analysis, we used administrative data and electronic health records for patients admitted to Boston Children's Hospital between March 1, 2012 and February 28, 2013. The table below describes the characteristics of patients with index hospitalizations during this time period (for patients for whom these data were available).

Table 4 – Boston Children's Hospital Cohort Patient Characteristics (Total N = 8,387)

| Patient Characteristic | | Number (%) of Index Hospitalizations |
|------------------------|---|--------------------------------------|
| Age | < 1 year | 2,113 (25.2) |
| | 1 to < 5 years | 2,041 (24.3) |
| | 5 to < 8 years | 978 (11.7) |
| | 8 to <12 years | 1,123 (13.4) |
| | 12 to < 18 years | 2,132 (25.4) |
| Gender | Female | 3,956 (47.2) |
| CCIs | CCI 1 - Infectious and parasitic disease | 11 (0.1) |
| | CCI 2 - Neoplasms | 272 (3.2) |
| | CCI 3 - Endocrine, nutritional, and metabolic diseases and immunity disorders | 873 (10.4) |
| | CCI 4 - Diseases of blood and blood-forming organs | 471 (5.6) |
| | CCI 5 - Mental disorders | 805 (9.6) |
| | CCI 6 - Diseases of the nervous system and sense organs | 1,234 (14.7) |
| | CCI 7 - Diseases of the circulatory system | 878 (10.5) |
| | CCI 8 - Diseases of the respiratory system | 1,042 (12.4) |
| | CCI 9 - Diseases of the digestive system | 557 (6.6) |
| | CCI 10 - Diseases of the genitourinary system | 157 (1.9) |
| | CCI 12 - Diseases of the skin and subcutaneous tissue | 69 (0.8) |
| | CCI 13 - Diseases of the musculoskeletal system | 337 (4.0) |
| | CCI 14 - Congenital anomalies | 2,212 (26.4) |
| | CCI 15 - Certain conditions originating in the perinatal period | 6 (0.1) |
| | CCI 16 - Symptoms, signs, and ill-defined conditions | 44 (0.5) |
| | CCI 17 - Injury and poisoning | 16 (0.2) |
| | CCI 18 - Factors influencing health status | 420 (5.0) |

| | | |
|----------------|----------------------------------|--------------|
| | and contact with health services | |
| CCI count | 0 or 1 body system | 5,858 (69.8) |
| | 2 body systems | 1,793 (21.4) |
| | 3 body systems | 602 (7.2) |
| | 4+ body systems | 134 (1.6) |
| Race/ethnicity | Asian/Pacific Islander | 290 (3.4) |
| | Black | 852 (10.2) |
| | Latino | 826 (9.8) |
| | Mixed | 15 (0.2) |
| | Native American | 731 (8.7) |
| | White | 5,054 (60.3) |
| | Missing | 618 (7.4) |

2a2. RELIABILITY TESTING

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter “see section 2b2 for validity testing of data elements”; and skip 2a2.3 and 2a2.4.

2a2.1. What level of reliability testing was conducted? (may be one or both levels)

☐ **Critical data elements used in the measure** (e.g., inter-abtractor reliability; data element reliability must address ALL critical data elements)

☒ **Performance measure score** (e.g., signal-to-noise analysis)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests

(describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used)

We evaluated the reliability of hospital-level readmission rates using the following formula:

$$\text{Reliability} = \sigma^2 / (\sigma^2 + V)$$

where σ^2 is the systematic variance among hospitals and V is the sampling variance of the sample estimate of a hospital's rate (both on the probability scale):

- $\sigma^2 = \sigma_L^2 \cdot p^2 \cdot (1-p)^2$
where σ_L^2 = variance component from model output in logit scale
- $V = p \cdot (1-p) / N$,
where p = the overall readmission rate across all hospitals and N = the hospital's volume

2a2.3. For each level checked above, what were the statistical results from reliability testing? (e.g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis)

Using the MAX dataset, we found that among the 1,743 total hospitals, 229 hospitals had a readmission rate reliability ≥ 0.5 ; these hospitals accounted for 62% of total LRI index hospitalizations. The readmission rate reliability was ≥ 0.7 for 70 hospitals, accounting for 35% of total LRI index hospitalizations.

Because hospitals with few pediatric patients would be less likely to participate in measuring pediatric readmissions, we evaluated readmission rate reliability for hospitals meeting selected minimum thresholds of pediatric all-condition and LRI index hospitalizations per year. We

determined that among the 539 hospitals with ≥ 100 annual index hospitalizations for any condition and ≥ 25 annual LRI index hospitalizations, readmission rate reliability was ≥ 0.5 for 229 hospitals, accounting for 74% of the LRI index hospitalizations at hospitals in this volume category. Readmission rate reliability was ≥ 0.7 for 70 hospitals, accounting for 42% of LRI index hospitalizations at hospitals in this volume category.

We found that among the 179 hospitals with ≥ 500 annual index hospitalizations and ≥ 25 annual LRI index hospitalizations, readmission rate reliability was ≥ 0.5 for 146 hospitals, accounting for 95% of the LRI index hospitalizations at hospitals in this volume category. Readmission rate reliability was ≥ 0.7 for 69 hospitals, accounting for 67% of LRI index hospitalizations at hospitals in this volume category.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.e., what do the results mean and what are the norms for the test conducted?)

Reliability values range from 0 to 1. If perfect information from a very large sample were available for a hospital, so that the hospital's random effect could be determined with perfect precision, then the reliability of that hospital's readmission rate would approach 1. If no information were available for a hospital, then the reliability of that hospital's readmission rate would be 0. Our results indicate that LRI readmission rates are reliable for hospitals accounting for a large proportion of index hospitalizations.

2b2. VALIDITY TESTING

2b2.1. What level of validity testing was conducted? (may be one or both levels)

- ☐ **Critical data elements** (data element validity must address ALL critical data elements)
- ☒ **Performance measure score**
 - ☒ **Empirical validity testing**
 - ☐ **Systematic assessment of face validity of performance measure score as an indicator of quality or resource use** (i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance)

2b2.2. For each level checked above, describe the method of validity testing and what it tests (describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)

Construct Validity

As detailed in Section 1a.2.1 of the Evidence form, many studies have provided evidence that readmission rates serve as a measure of healthcare quality. Use of approaches to diagnosis, treatment, and monitoring of disease that adhere to clinical practice guidelines has been correlated with lower readmission rates.^{1–3} Likewise, improvements in the quality of clinical management have been associated with reductions in readmissions.^{4–6} Readmission rates have also been found to reflect the quality of discharge and care transition processes. Several studies, mostly in adults, have demonstrated that interventions focused on improving these processes have been linked with decreased readmissions, suggesting that the quality of these processes is associated with readmission risk.^{7–26}

Although the medical literature provides ample evidence for the relationship between quality of care and pediatric and adult readmission risk, assessing the construct validity of pediatric readmission rates directly by examining the correlation of rates with other pediatric measures of quality does not appear to be currently feasible. To perform this analysis, pediatric inpatient

claims-based quality measures or large, multi-hospital datasets of scores from pediatric quality measures would be required. However, to our knowledge, no other publicly available claims-based pediatric inpatient quality measures exist, including among the Healthcare Effectiveness Data and Information Set (HEDIS) measures, the CHIPRA Initial Core Set of Children's Health Care Quality Measures, and other measure collections that we examined. There also do not appear to be large datasets with scores from pediatric inpatient quality measures.

Criterion Validity

We evaluated the ability of our measure to identify the outcome of interest, readmission, from administrative data by comparing the measure's performance against the gold standard of chart reviews. We performed this analysis using administrative data and electronic health records for patients admitted to Boston Children's Hospital over a 1-year period (see Section 1.7 above for a summary of patient characteristics). We determined from the administrative data that 8,833 index hospitalizations occurred during this time period. We then identified hospitalizations that met measure criteria for readmissions (i.e., the readmissions were not for a planned procedure or chemotherapy) in 2 ways: (1) analysis of the administrative data using the measure program and (2) review of electronic health records. We assessed the health records by first examining whether each index hospitalization was followed by a readmission within 30 days based on presence of inpatient admission orders (such an order is entered for every hospitalization). We then reviewed clinical documentation, including admission notes, discharge summaries, and procedure notes, for 500 randomly selected readmissions to determine whether the readmission had been for a planned procedure or chemotherapy.

Validity of Planned Procedure Algorithm

We also verified the face validity of the planned procedure algorithm used to identify hospitalizations for planned procedures. We sought public comments on the algorithm in a Federal Register Notice.²⁷ No comments were submitted to suggest that procedures be removed from the list of planned procedures because they are not typically planned or are not a reason for hospitalization. Twenty-four procedures were submitted with the suggestion that they be added to the list of planned procedures. Of these, 7 were already included on the planned procedure list; our expert clinicians in 3 relevant specialties reviewed the remaining 17 procedures. Most of the remaining codes were for procedures for which patients are not hospitalized. Based on the experts' review, however, 2 procedures, both organ transplantation procedures, were added to the planned procedure list. These transplantation procedures originally had been excluded because they did not meet our operational definition of "planned" (i.e., scheduled at least 24 hours in advance), but it was agreed that they should be added as a special case of procedures for which the need is typically known in advance, even though the actual operation occurs urgently once an organ becomes available. For the same reason, 9 other transplantation procedures were also added to the planned procedure list.

Identification of International Classification of Diseases, 10th Revision (ICD-10) Codes

To identify ICD-10 codes for chronic conditions, mental health conditions, and obstetric conditions, we used AHRQ's ICD-10 version of its Chronic Condition Indicator tool. For all other codes used in the measure, we obtained ICD-10 codes by performing conversions from the ICD-9 codes we had selected during measure development. Our goal for the conversions was to compile an ICD-10 code set that was fully consistent with the intent of the original ICD-9 set. We carried out the conversions using the 3M™ Code Translation Tool and reviewed all conversions to ensure that the resulting ICD-10 codes captured the intended concepts, removing ICD-10 codes from the code set as appropriate.

2b2.3. What were the statistical results from validity testing? (e.g., correlation; t-test)

The sensitivity and specificity of the measure for identifying eligible readmissions from administrative data were 87.0% and 99.7%, respectively.

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i.e., what do the results mean and what are the norms for the test conducted?)

The measure is able to identify eligible readmissions from administrative data with high sensitivity and specificity.^{28–31} We found face validity for our planned procedure algorithm.

2b3. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — skip to section 2b4

2b3.1. Describe the method of testing exclusions and what it tests (describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used)

We chose to exclude hospitalizations with a primary mental health diagnosis from the measure cohort after evaluating the relationship between the primary diagnosis and the readmission outcome. We fitted a hierarchical random slopes regression model to the data. The model consisted of patients nested within hospitals at the first level and 2 random slope indicator variables at the second level: (a) an indicator variable for the primary diagnosis of interest alone and (b) an indicator variable for all other possible primary diagnoses.

2b3.2. What were the statistical results from testing exclusions? (include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores)

We excluded 1,211,196 total hospitalizations from the measure. Of these, 25,459 (2.1%) were excluded for a primary diagnosis of a mental health condition. The median hospital percentage of index hospitalizations with a primary diagnosis of a mental health condition was 0.0% (IQR 0.0%–1.6%). In the analysis described in Section 2b3.1, we found that for primary diagnoses other than mental health conditions, the regression coefficient for the primary diagnosis of interest had a positive correlation with the regression coefficient for all other diagnoses, suggesting that performance on readmissions for the primary diagnosis of interest tends to correspond with performance on readmissions for all other diagnoses (the converse is also true). However, the regression coefficient for primary mental health diagnoses had a negative correlation with the coefficient for non-mental health diagnoses, suggesting that performance on readmissions for mental health conditions does not tend to correspond with performance on readmissions for non-mental health conditions.

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (i.e., the value outweighs the burden of increased data collection and analysis. *Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

Hospitalizations with a primary diagnosis code for a mental health condition are excluded from the measure cohort because we found that hospitals with high readmission rates for mental health hospitalizations tend to have low readmission rates for hospitalizations for other conditions, and vice versa.

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used?

- ☐ No risk adjustment or stratification
- ☒ Statistical risk model with **20 fixed effect variables representing 4 types of** risk factors
- ☐ Stratification by [Click here to enter number of categories](#) risk categories
- ☐ Other, [Click here to enter description](#)

2b4.2. If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors used in the statistical risk model or for stratification by risk (e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of $p < 0.10$; correlation of x or higher; patient factors should be present at the start of care and not related to disparities)

Identification of Candidate Variables for the Case-Mix Adjustment Model

To choose candidate variables for possible inclusion in the case-mix adjustment model, we considered patient demographic and clinical characteristics based on review of readmission studies and other readmission measures. We selected as candidate variables patient characteristics that (1) may differ in their distribution across hospitals, (2) may be related to readmission risk, and (3) are less likely to be related to variation in pediatric quality across health systems. Adjusting for such variables accounts for the variation in readmission rates that is attributable to differences in the distribution of patient characteristics across hospitals. In deciding on variables for the final model, we also took into account the variables and variable specifications used in the case-mix adjustment model for our Pediatric All-Condition Readmission Model, thus ensuring measure harmonization.

We examined the following candidate case-mix variables for use in the measure model: age, gender, presence of chronic conditions in 18 different body systems, and number of body systems affected by chronic conditions.

Methods for Variable Selection

Step 1: Determine Bivariate Relationship with Readmission and Select Parsimonious Specifications of Candidate Variables

We assessed the relationship between each candidate variable and readmission risk in bivariate analysis. The analysis included the candidate variable as well as a hospital random intercept. We excluded variables from further analysis if no specification for the candidate variable was found to have a statistically significant relationship with readmission. Throughout the variable selection process, we used a p-value $< .05$ as the criterion for statistical significance.

For candidate variables that were significantly related to readmission and could be specified in > 1 way (e.g., age can be expressed as a continuous or categorical variable), we evaluated multiple potential variable specifications. Candidate variables that could be specified either continuously or categorically were created using both approaches to assess whether a linear specification provided the best fit. For variables that were specified both continuously and categorically, and for differing specifications of cut-points for the levels of categorical variables, we determined the best-fitting specification of each variable based on the likelihood ratio chi-

square values. If 2 specifications had close likelihood ratios, we chose which specification of the variable to evaluate further in multivariate analysis based on parsimony and clinical face validity.

Step 2: Determine Statistical Significance of Candidate Variables in Multivariate Analysis

We used a multivariate model to assess whether variables that were statistically significant in bivariate analysis remained significant in multivariate analysis.

2b4.4. What were the statistical results of the analyses used to select risk factors?

Age

We found that age had a non-linear statistically significant relationship with 30-day readmission in both bivariate and multivariate analysis (see Table 5). The final specification for age is detailed below. We chose the specification because (a) the categorical variable captures the non-linear relationship of age with the outcome of readmission, (b) the specification has a high likelihood ratio chi-square relative to other less parsimonious specifications, and (c) the age group categories are clinically and developmentally meaningful.

Table 5 – Bivariate and Multivariate Results for Age

| Age (agegroup) | Bivariate Analysis | | Multivariate Analysis | |
|-------------------------|---------------------------|---------|------------------------------|---------|
| | OR | p-value | OR | p-value |
| Likelihood ratio 139.70 | | | | |
| 1 = 0 ≤ age < 1 | reference | - | reference | - |
| 2 = 1 ≤ age < 5 | 0.70 | < .001 | 0.62 | < .001 |
| 3 = 5 ≤ age < 8 | 0.55 | < .001 | 0.42 | < .001 |
| 4 = 8 ≤ age < 12 | 0.74 | .001 | 0.45 | < .001 |
| 5 = 12 ≤ age < 18 | 1.10 | .26 | 0.59 | < .001 |

Gender

We found that male gender was significantly associated with increased odds of readmission in bivariate and multivariate analysis (see Table 6) and therefore retained the variable in our model.

Table 6 – Bivariate and Multivariate Results for Gender

| Gender (male) | Bivariate Analysis | | Multivariate Analysis | |
|------------------------|---------------------------|---------|------------------------------|---------|
| | OR | p-value | OR | p-value |
| Likelihood ratio 19.72 | | | | |
| 0 = female | reference | - | reference | - |
| 1 = male | 1.15 | < .001 | 1.14 | < .001 |

Chronic Conditions

To account for chronic disease comorbidity, we used the AHRQ CCI tool to classify ICD-9-CM diagnosis codes for chronic conditions into 18 body systems (organ systems, disease categories, or other categories). We created a dichotomous variable for each body system, with a value of 1 if ≥ 1 chronic condition was present (coded as a primary or secondary diagnosis for an index hospitalization) in that body system or 0 if no chronic condition was present in that body system. We examined each of the 18 CCI variables in relation to the outcome of readmission in bivariate and multivariate analysis, using absence of a chronic condition in the body system in question as the reference.

We found that 12 of the 18 dichotomous variables were significantly related to readmission in bivariate analysis and 11 of the 18 were significantly related in multivariate analysis.

We chose to retain all but the CCI for body system 11, “Complications of pregnancy, childbirth, and the puerperium,” in the final model (a) to maintain, to the extent possible, the coherence of the complete AHRQ CCI tool, (b) because most of the CCI variables had a statistically significant relationship with the outcome of readmission, and (c) for harmonization with our Pediatric All-Condition Readmission Measure. Patients who have a primary diagnosis code for an obstetric condition or any diagnosis or procedure code for delivery are excluded from the measure cohort. We have found using various datasets that this exclusion leaves very few (or sometimes no) patients who have a secondary diagnosis code for a chronic condition within body system 11, which could create model-fitting problems if CCI 11 were included in the case-mix-adjustment model.

Table 7 — Bivariate and Multivariate Results for CCIs

| CCI (cci) | | Bivariate | | Multivariate | |
|---|---|-----------|---------|--------------|---------|
| Likelihood ratio range: 0.19 (cci1) to 340.07 (cci14) | | OR | p-value | OR | p-value |
| 1 | Infectious and parasitic disease | 1.41 | .64 | 1.19 | 0.82 |
| 2 | Neoplasms | 2.26 | < .001 | 2.75 | < .001 |
| 3 | Endocrine, nutritional, and metabolic diseases and immunity disorders | 2.34 | < .001 | 2.19 | < .001 |
| 4 | Diseases of blood and blood-forming organs | 1.40 | < .001 | 1.66 | < .001 |
| 5 | Mental disorders | 1.95 | < .001 | 1.51 | < .001 |
| 6 | Diseases of the nervous system and sense organs | 3.00 | < .001 | 2.49 | < .001 |
| 7 | Diseases of the circulatory system | 2.49 | < .001 | 1.79 | < .001 |
| 8 | Diseases of the respiratory system | 0.92 | .03 | 1.18 | < .001 |
| 9 | Diseases of the digestive system | 2.55 | < .001 | 2.05 | < .001 |
| 10 | Diseases of the genitourinary system | 3.79 | < .001 | 2.43 | < .001 |
| 11 | Complications of pregnancy, childbirth, and the puerperium | No cases | | | |
| 12 | Diseases of the skin and subcutaneous tissue | 1.53 | .09 | 1.44 | .17 |
| 13 | Diseases of the musculoskeletal system | 1.69 | .01 | 1.08 | .76 |
| 14 | Congenital anomalies | 2.82 | < .001 | 2.23 | < .001 |
| 15 | Certain conditions originating in the perinatal period | 1.72 | .38 | 1.24 | .73 |
| 16 | Symptoms, signs, and ill-defined conditions | 0.43 | .15 | 0.34 | .07 |
| 17 | Injury and poisoning | 3.61 | .25 | 1.70 | .67 |
| 18 | Factors influencing health status and contact with health services | 3.50 | < .001 | 2.51 | < .001 |

Number of body systems affected by chronic conditions

We also evaluated a count variable of the number of body systems in which a chronic condition was present for each index hospitalization. To avoid problems with model estimation, we top-coded the variable at ≥ 4 systems affected by a chronic disease because there were few index admissions with diagnoses in ≥ 5 systems.

The CCI count variable had a statistically significant relationship with readmission in bivariate and multivariate analysis. In multivariate analysis, an increasing count was associated with decreasing odds of readmission. This finding indicates that the presence of chronic conditions in multiple body systems confers a readmission risk that is lower than the sum of the risk of chronic conditions in each of the individual body systems, such that the CCI count variable

serves to prevent overestimation of the risk associated with having chronic conditions in multiple body systems. We chose to retain *cci count* as an ordinal variable based on (a) its statistically significant relationship with the outcome of readmission and (b) because it adjusts the readmission risk associated with having chronic conditions in multiple body systems.

Table 8 – Bivariate and Multivariate Results for CCI Count

| CCI Count (<i>cci count</i>) | Bivariate | | Multivariate | |
|--------------------------------|-----------|---------|--------------|---------|
| | OR | p-value | OR | p-value |
| Likelihood ratio 547.89 | | | | |
| 1 = 0 to 1 | reference | - | reference | - |
| 2 = 2 | 2.46 | < .001 | 0.93 | < .001 |
| 3 = 3 | 3.65 | < .001 | 0.74 | < .001 |
| 4 = 4 or more | 4.45 | < .001 | 0.40 | < .001 |

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

We assessed the discriminative ability of the model using the c-statistic.^{32,33} Discrimination refers to how well the model distinguishes between subjects with and without the outcome (in this case, readmission).³² The c-statistic is a unitless measure of the probability that a randomly selected subject who experienced readmission will have a higher predicted probability of having been readmitted than a randomly selected subject who did not experience readmission.³²

We assessed model calibration with a chi-square goodness-of-fit test analogous to the Hosmer-Lemeshow test.³⁴ We used the test, which evaluates how well observed outcomes correspond to those predicted by the fitted logistic regression model,³⁴ to determine how well observed and predicted numbers of readmissions matched for the levels of the 2 ordinal variables in our case-mix adjustment model, age and CCI count. The lack of a significant difference between observed and predicted values indicates good model calibration

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

if stratified, skip to 2b4.9

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*):

The c-statistic for our case-mix adjustment model, when applied to the MAX dataset, was 0.71.

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

When we stratified records by age categories, the p-value for the chi-square goodness-of-fit test was not significant (p = .56).

Table 9 – Chi-Square Goodness-of-Fit Test for Lower Respiratory Infection Readmissions: Age

| Age Group | Number of Index Admissions | Predicted Cases of Readmissions | Observed Cases of Readmissions |
|-----------|----------------------------|---------------------------------|--------------------------------|
| | | n (%) | n (%) |
| 0 years | 33,149 | 2,105 (6.4%) | 2,152 (6.5%) |
| 1-4 years | 24,710 | 1,154 (4.7%) | 1,182 (4.8%) |
| 5-7 years | 4,501 | 167 (3.7%) | 171 (3.8%) |

| | | | |
|-------------|-------|------------|------------|
| 8-11 years | 2,544 | 127 (5.0%) | 130 (5.1%) |
| 12-17 years | 2,256 | 170 (7.5%) | 173 (7.7%) |

When we stratified records by categories of the number of body systems affected by chronic conditions, the p-value for the chi-square goodness-of-fit test also was not significant (p = .32).

Table 10 – Chi-Square Goodness-of-Fit Test for Lower Respiratory Infection Readmissions: CCI Count

| CCI Count | Number of Index Admissions | Predicted Cases of Readmissions | Observed Cases of Readmissions |
|---------------------|----------------------------|---------------------------------|--------------------------------|
| | | n (%) | n (%) |
| 0 or 1 body systems | 60,546 | 2,881 (4.8%) | 2,959 (4.9%) |
| 2 body systems | 4,203 | 435 (10.3%) | 440 (10.5%) |
| 3 body systems | 1,567 | 252 (16.1%) | 254 (16.2%) |
| 4+ body systems | 844 | 154 (18.3%) | 155 (18.4%) |

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Figure 1 – Chi-Square Goodness-of-Fit Test for Lower Respiratory Infection Readmissions: Age

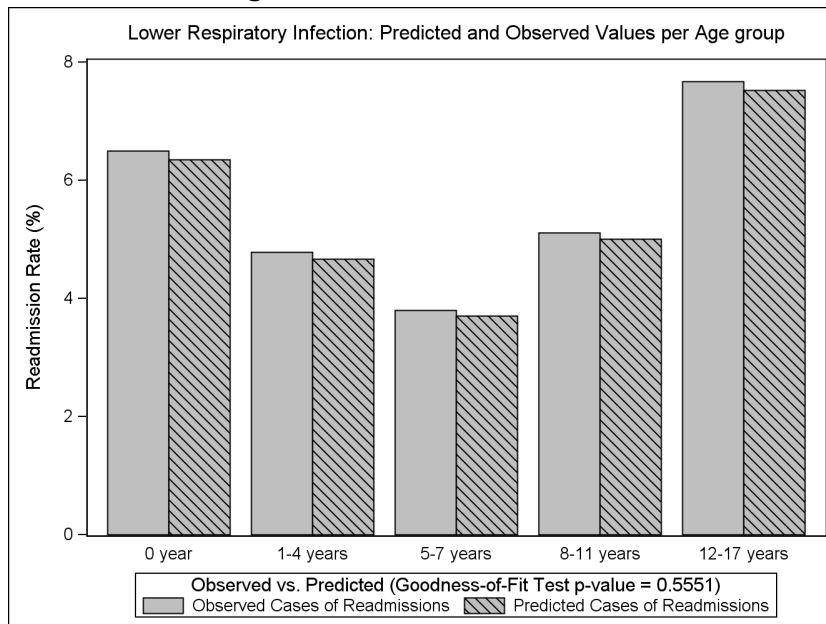
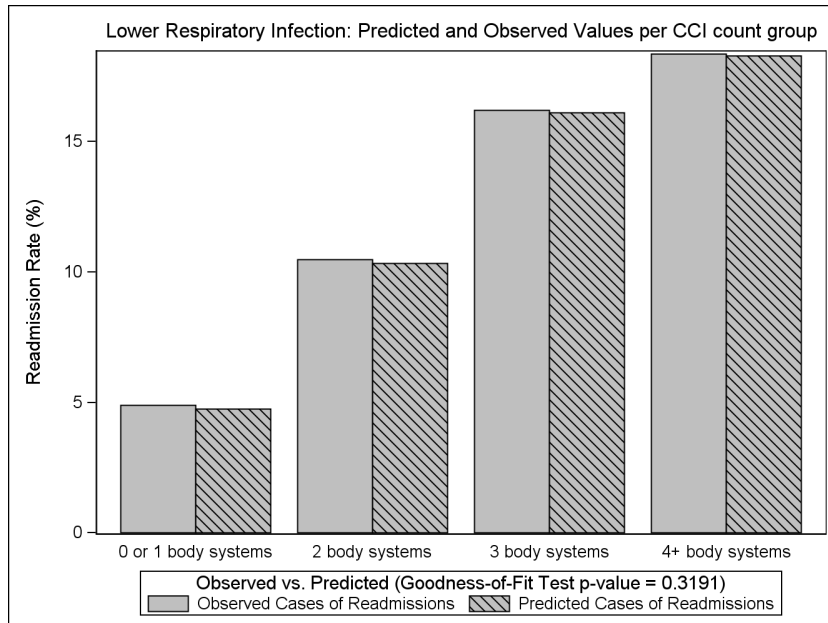


Figure 2 — Chi-Square Goodness-of-Fit Test for Lower Respiratory Infection Readmissions: CCI Count



2b4.9. Results of Risk Stratification Analysis:

Not applicable.

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.e., what do the results mean and what are the norms for the test conducted)

The discriminative ability of the case-mix adjustment model is good, with a c-statistic that is very similar to that of other 30-day readmission measures.^{28–31} The model calibration is also good, with a close match between observed and predicted numbers of readmissions.

***2b4.11. Optional Additional Testing for Risk Adjustment** (not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods)

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)

We identified hospitals with meaningfully different readmission performance based on their excess readmission ratio, calculated using NQF-endorsed methods.³⁵ For each hospital, the numerator of the ratio, its number of adjusted actual readmissions, is calculated by estimating the probability of readmission for each patient at that hospital and adding the probabilities for all of the hospital's patients. The denominator of the ratio, its number of expected readmissions, is calculated by estimating the probability of readmission for each of the hospital's patients if he or she had been at an average hospital and then adding the probabilities for all of the hospital's patients.

Numerator — Adjusted Actual Readmissions

$$\text{Each patient's predicted probability of readmission} = \frac{1}{1 + e^{-Z_a}}$$

Z_a = hospital-specific effect + $X\beta$

where $X\beta$ = intercept + case-mix adjustment coefficients

Denominator — Expected Readmissions

$$\text{Each patient's predicted probability of readmission} = \frac{1}{1 + e^{-Z_e}}$$

$Z_e = X\beta$

where $X\beta$ = intercept + case-mix adjustment coefficients

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined)

Of 1,743 hospitals, 561 (32%) had an excess readmission ratio > 1, indicating that their number of adjusted actual readmissions was higher than would be expected at an average hospital. Among hospitals with an excess readmission ratio > 1, the median ratio was 1.18 (IQR 1.09–1.33). In other words, for half of the hospitals with an excess readmission ratio > 1, their number of adjusted actual readmissions exceeded their number of expected readmissions by > 18%.

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.e., what do the results mean in terms of statistical and meaningful differences?)

The measure can identify hospitals with meaningfully different readmission performance.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
If only one set of specifications, this section can be skipped.

Note: This criterion is directed to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator). **If comparability is not demonstrated, the different specifications should be submitted as separate measures.**

2b6.1. Describe the method of testing conducted to demonstrate comparability of performance scores for the same entities across the different datasources/specifications (describe the steps—do not just name a method; what statistical analysis was used)

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (e.g., correlation, rank order)

2b6.3. What is your interpretation of the results in terms of demonstrating comparability of performance measure scores for the same entities across the different data sources/specifications? (i.e., what do the results mean and what are the norms for the test conducted)

References

1. Bozic KJ, Maselli J, Pekow PS, Lindenauer PK, Vail TP, Auerbach AD. The influence of procedure volumes and standardization of care on quality and efficiency in total joint replacement surgery. *J Bone Joint Surg Am*. 2010;92(16):2643–2652.
2. Ludke RL, MacDowell NM, Booth BM, Hunter SA. Appropriateness of admissions and discharges among readmitted patients. *Health Serv Res*. 1990;25(3):501–525.
3. Heidenreich PA, Hernandez AF, Yancy CW, Liang L, Peterson ED, Fonarow GC. Get With The Guidelines program participation, process of care, and outcome for Medicare patients hospitalized with heart failure. *Circ Cardiovasc Qual Outcomes*. 2012;5(1):37–43. d
4. Fassel BA, Nkoy FL, Stone BL, Srivastava R, Simon TD, Uchida DA, Koopmeiners K, Greene T, Cook LJ, Maloney CG. The Joint Commission Children's Asthma Care quality measures and asthma readmissions. *Pediatrics*. 2012;130(3):482–491.
5. Pillai D, Song X, Pastor W, Ottolini M, Powell D, Wiedermann BL, DeBiasi RL. Implementation and impact of a consensus diagnostic and management algorithm for complicated pneumonia in children. *J Invest Med*. 2011;59(8):1221–1227.
6. Cheney J, Barber S, Altamirano L, Medico Cirujano, Cheney M, Williams C, Jackson M, Yates P, O'Rourke P, Wainwright C. A clinical pathway for bronchiolitis is effective in reducing readmission rates. *J Pediatr*. 2005;147(5):622–626.
7. Inglis SC, Clark RA, McAlister FA, Ball J, Lewinter C, Cullington D, Stewart S, Cleland JG. Structured telephone support or telemonitoring programmes for patients with chronic heart failure. *Cochrane Database Syst Rev Online*. 2010;(8):CD007228.
8. Phillips CO, Wright SM, Kern DE, Singa RM, Shepperd S, Rubin HR. Comprehensive discharge planning with postdischarge support for older patients with congestive heart failure: a meta-analysis. *JAMA*. 2004;291(11):1358–1367. doi:10.1001/jama.291.11.1358.
9. Scott IA. Preventing the rebound: improving care transition in hospital discharge processes. *Aust Health Rev*. 2010;34(4):445–451. doi:10.1071/AH09777.
10. Shepperd S, Lannin NA, Clemson LM, McCluskey A, Cameron ID, Barras SL. Discharge planning from hospital to home. *Cochrane Database Syst Rev Online*. 2013;1:CD000313.
11. Sochalski J, Jaarsma T, Krumholz HM, Laramée A, McMurray JJV, Naylor MD, Rich MW, Riegel B, Stewart S. What works in chronic care management: the case of heart failure. *Health Aff (Millwood)*. 2009;28(1):179–189.
12. Anderson C, Deepak BV, Amoateng-Adjepong Y, Zarich S. Benefits of comprehensive inpatient education and discharge planning combined with outpatient support in elderly patients with congestive heart failure. *Congest Heart Fail*. 2005;11(6):315–321.
13. Balaban RB, Weissman JS, Samuel PA, Woolhandler S. Redefining and redesigning hospital discharge to enhance patient care: a randomized controlled study. *J Gen Intern Med*. 2008;23(8):1228–1233.
14. Coleman EA, Parry C, Chalmers S, Min S-J. The care transitions intervention: results of a randomized controlled trial. *Arch Intern Med*. 2006;166(17):1822–1828.
15. Evans RL, Hendricks RD. Evaluating hospital discharge planning: a randomized clinical trial. *Med Care*. 1993;31(4):358–370.
16. Jack BW, Chetty VK, Anthony D, Greenwald JL, Sanchez GM, Johnson AE, Forsythe SR, O'Donnell JK, Paasche-Orlow MK, Manasseh C, Martin S, Culpepper L. A reengineered hospital discharge program to decrease rehospitalization: a randomized trial. *Ann Intern Med*. 2009;150(3):178–187.
17. Naylor MD, Brooten DA, Campbell RL, Maislin G, McCauley KM, Schwartz JS. Transitional care of older adults hospitalized with heart failure: a randomized, controlled trial. *J Am Geriatr Soc*. 2004;52(5):675–684.
18. Naylor M, Brooten D, Jones R, Lavizzo-Mourey R, Mezey M, Pauly M. Comprehensive discharge planning for the hospitalized elderly. A randomized clinical trial. *Ann Intern Med*. 1994;120(12):999–1006.

19. Koehler BE, Richter KM, Youngblood L, Cohen BA, Prengler ID, Cheng D, Masica AL. Reduction of 30-day postdischarge hospital readmission or emergency department (ED) visit rates in high-risk elderly medical patients through delivery of a targeted care bundle. *J Hosp Med*. 2009;4(4):211–218.
20. Rich MW, Beckham V, Wittenberg C, Leven CL, Freedland KE, Carney RM. A multidisciplinary intervention to prevent the readmission of elderly patients with congestive heart failure. *N Engl J Med*. 1995;333(18):1190–1195.
21. Costantino ME, Frey B, Hall B, Painter P. The influence of a postdischarge intervention on reducing hospital readmissions in a Medicare population. *Popul Heal Manag*. 2013;16(5):310–316.
22. Flood KL, Maclennan PA, McGrew D, Green D, Dodd C, Brown CJ. Effects of an acute care for elders unit on costs and 30-day readmissions. *JAMA Intern Med*. 2013:1–7.
23. Hoffmann RG, Yan K, Brousseau DC. Outpatient follow-up and rehospitalizations for sickle cell disease patients. *Pediatr Blood Cancer*. 2012;58(3):406–409.
24. Bradley EH, Curry L, Horwitz LI, Sipsma H, Wang Y, Walsh MN, Goldmann D, White N, Piña IL, Krumholz HM. Hospital strategies associated with 30-day readmission rates for patients with heart failure. *Circ Cardiovasc Qual Outcomes*. 2013;6(4):444–450.
25. Harrison PL, Hara PA, Pope JE, Young MC, Rula EY. The impact of postdischarge telephonic follow-up on hospital readmissions. *Popul Heal Manag*. 2011;14(1):27–32.
26. Kripalani S, Theobald CN, Anctil B, Vasilevskis EE. Reducing Hospital Readmission Rates: Current Strategies and Future Directions. *Annu Rev Med*. 2013.
27. Department of Health and Human Services, Agency for Healthcare Research and Quality. Request for comments on pediatric planned procedure algorithm. *Fed Regist*. 2013;78:57639 –57640.
28. Horwitz L, Partovian C, Lin Z, Herrin J, Grady J, Conover M, Montague J, Dillaway C, Bartczak K, Suter L, Ross J, Bernheim S, Krumholz H, Drye E. *Hospital-Wide All-Cause Unplanned Readmission Measure: Final Technical Report*.; 2012.
29. Grosso LM, Curtis JP, Lin Z, Geary LL, Vellanky S, Oladele C, Ott LS, Parzynski C, Bernheim S, Suter LG, Drye EE, Krumholz HM. *Hospital-level 30-Day All-Cause Risk-Standardized Readmission Rate Following Elective Primary Total Hip Arthroplasty (THA) And/Or Total Knee Arthroplasty (TKA): Measure Methodology Report*.; 2012.
30. Grady JN, Lin Z, Wang C, Keenan M, Nwosu C, Bhat KR, Horwitz LI, Drye EE, Krumholz HM, Bernheim SM. *2013 Measures Updates and Specifications Report: Hospital-Level 30-Day Risk-Standardized Readmission Measures for Acute Myocardial Infarction, Heart Failure, and Pneumonia (Version 6.0)*; 2013.
31. Rice-Townsend S, Hall M, Barnes JN, Lipsitz S, Rangel SJ. Variation in risk-adjusted hospital readmission after treatment of appendicitis at 38 children's hospitals: an opportunity for collaborative quality improvement. *Ann Surg*. 2013;257(4):758–765.
32. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. *BMC Med Res Methodol*. 2012;12:82.
33. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128–138.
34. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med*. 1997;16(9):965–980.
35. Department of Health and Human Services. Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Fiscal Year 2013 Rates; Hospitals' Resident Caps for Graduate Medical Education Payment Purposes; Quality Reporting Requirements for Specific Providers and for Ambulatory Surgical Centers. *Fed Regist*. 2012;77(170).