NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Click to go to the link. ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 3479

Corresponding Measures:

De.2. Measure Title: Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRF)

Co.1.1. Measure Steward: Centers for Medicare & Medicaid Services

De.3. Brief Description of Measure: The Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (DTC-PAC IRF) was developed to address the resource use and other measures domain of Discharge to the Community mandated by the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This outcome measure assesses successful discharge to community from an IRF, with successful discharge to community including no unplanned rehospitalizations and no death in the 31 days following IRF discharge. The measure reports an IRF's risk-standardized rate of Medicare fee-for-service (FFS) patients who are discharged to the community following an IRF stay, and do not have an unplanned readmission to an acute care hospital or long-term care hospital (LTCH) in the 31 days following discharge to community, and who remain alive during the 31 days following discharge to community. The measure is calculated using two consecutive years of Medicare FFS claims data and was developed using calendar year (CY) 2012-2013 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., IRF discharges from October 1, 2015 through September 30, 2017.

The measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the IRF Quality Reporting Program (QRP) finalized in the FY 2017 IRF Prospective Payment System (PPS) Final Rule and implementation began October 1, 2016 [1]. Confidential feedback reports on measure performance were distributed to IRF providers in Fall 2017. The measure will be publicly reported on the IRF Compare website (https://www.medicare.gov/inpatientrehabilitationfacilitycompare/) in Fall 2018 using FY 2016-2017 data. Four claims-based discharge to community measures were developed for IRF, LTCH, skilled nursing facility, and home health agency settings to meet the mandate of the IMPACT Act. These measures were conceptualized uniformly across the four settings, in terms of the definition of the discharge to community outcome, the approach to risk-adjustment, and the measure calculation.

References

[1] Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 Federal Register, Vol. 81, No. 151. https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf

1b.1. Developer Rationale: Successful discharge to community from IRF is widely recognized as an important outcome for patients and their families and an indicator of IRF quality of care.[1-6] The Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act of 2014) mandated the specification of a measure to address the domain of discharge to community across post-acute care (PAC) settings.[4] Discharge to community is an important goal for both patients with potential for functional improvement, and patients who may not be expected to make functional gains given their clinical condition or disease. IRFs frequently track their discharge to community outcomes for internal quality improvement purposes; however, this tracking commonly assesses discharge destination alone and not post-discharge readmissions or death. By assessing whether patients successfully remain in the community without acute complications for 31 days following discharge, the DTC-PAC IRF measure is a more meaningful, patient- and family-centered measure of successful community discharge. Discharge to community is an actionable health care outcome as targeted interventions have been shown to successfully increase discharge to community rates in a variety of post-acute settings, including IRFs [7,8] (see attached evidence form). The effectiveness of these interventions suggests that improvement in discharge to community rates among IRF patients is possible through modifying provider-led processes and interventions. In addition to being important from a patient and family perspective, avoiding institutionalization and promoting discharge to community when appropriate are also important from a cost and resource use perspective. Patients discharged to community settings, on average, incur lower costs over the recovery episode, compared with patients discharged to institutional settings.[9-11]

Measuring and publicly reporting the DTC-PAC IRF measure is expected to help differentiate IRFs based on quality of care, and drive improvement in this outcome. The DTC-PAC IRF measure would impact a large number of Medicare beneficiaries and the overall Medicare program. In 2016, Medicare spent $7.7 billion on fee-for service (FFS) IRF care; this care was provided in about 1,200 IRFs nationwide, covering more than 391,000 IRF stays for about 350,000 beneficiaries.[5] There are no NQF-endorsed measures assessing successful discharge to community from IRFs.

[1] Gillsjö C, Schwartz-Barcott D, von Post I. Home: The place the older adult can not imagine living without. BMC Geriatrics. 2011;11(1):10.

[2] Unsworth C. Clients´ perceptions of discharge housing decisions after stroke rehabilitation. American Journal of Occupational Therapy. 1996;50(3):207-216.

[3] Roush CV, Cox JE. The meaning of home: how it shapes the practice of home and hospice care. Home Healthcare Now. 2000;18(6):388-394.

[4] Improving Medicare Post-Acute Care Transformation Act of 2014. Public Law 113–185—October 6, 2014. Available at: https://www.gpo.gov/fdsys/pkg/PLAW-113publ185/pdf/PLAW-113publ185.pdf.

[5] Medicare Payment Advisory Commission. Report to the Congress: Medicare Payment Policy. March 2018. Available at: http://www.medpac.gov/docs/default-source/reports/mar18_medpac_entirereport_sec.pdf.

[6] Hsieh C-H, DeJong G, Groah S, Ballard PH, Horn SD, Tian W. Comparing rehabilitation services and outcomes between older and younger people with spinal cord injury. Archives of Physical Medicine and Rehabilitation. 2013;94(4, Supplement):S175-S186.

[7] Kushner DS, Peters KM, Johnson-Greene D. Evaluating Siebens Domain Management Model for inpatient rehabilitation to increase functional independence and discharge rate to home in geriatric patients. Archives of Physical Medicine and Rehabilitation. 2015;96(7):1310-1318.

[8] Kushner DS, Peters KM, Johnson-Greene D. Evaluating use of the Siebens Domain Management Model during inpatient rehabilitation to increase functional independence and discharge rate to home in stroke patients. PM&R: the Journal of Injury, Function, and Rehabilitation. 2015;7(4):354-364.

[9] Dobrez D, Heinemann AW, Deutsch A, Manheim L, Mallinson T. Impact of Medicare´s prospective payment system for inpatient rehabilitation facilities on stroke patient outcomes. American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists. 2010;89(3):198-204.

[10] Gage B, Morley M, Spain P, Ingber M. Examining Post Acute Care Relationships in an Integrated Hospital System Final Report. RTI International; 2009.

[11] Medicare Payment Advisory Commission. Paying for sequential stays in a unified prospective payment system for post-acute care. In: Report to the Congress: Medicare and the Health Care Delivery System. June 2018. Available at: http://www.medpac.gov/docs/default-source/reports/jun18_ch4_medpacreport_sec.pdf?sfvrsn=0

S.4. Numerator Statement: The measure numerator is the risk-adjusted predicted estimate of the number of patients who are discharged to the community, and do not have an unplanned readmission to an acute care hospital or LTCH in the 31-day post-discharge observation window, and who remain alive during the post-discharge observation window.

This estimate starts with the observed number of discharges to community, defined as:

(i) discharges to home or self care with or without home health services, based on Patient Discharge Status Codes 01, 06, 81, or 86 on the Medicare FFS claim [2]; and

(ii) no unplanned acute or LTCH hospitalizations in the 31-day post-discharge window; and

(iii) no death in the 31-day post-discharge window.

The discharge to community outcome is risk-adjusted for patient characteristics and a statistical estimate of the facility effect beyond case-mix (described below).

References

[2] National Uniform Billing Committee Official UB-04 Data Specifications Manual 2018, Version 12, July 2017, Copyright 2017, American Hospital Association.

S.6. Denominator Statement: The target population for the measure is the group of Medicare FFS beneficiaries who are discharged from an IRF during the measurement period and are not excluded based on the measure exclusion criteria (see S.8. and S.9.).

The measure denominator is the risk-adjusted expected number of discharges to community. This estimate includes risk-adjustment for patient characteristics with the facility effect removed. The “expected” number of discharges to community is the predicted number of risk-adjusted discharges to community if the same patients were treated at the average facility. The logistic regression model used to calculate the denominator is developed using all non-excluded facility stays in the national data. The denominator is computed in the same way as the numerator, but the facility effect is set at the average.

S.8. Denominator Exclusions: Measure exclusion criteria are based on administrative data from Medicare claims and eligibility files. Exclusion criteria were selected to maintain clinical validity of the measure by excluding stays for which discharge to community would not be appropriate, to ensure data availability and completeness, to exclude stays with problematic claims data, and to maintain relevance to the IRF QRP (e.g., excluding IRFs not included in the IRF QRP based on regional location). Only IRF stays that are preceded by a short-term acute care stay in the 30 days prior to the IRF admission date are included in the measure; this is because risk-adjustment variables come from the short-term acute care stay in the 30 days prior to IRF admission. Stays ending in transfers to the same level of care (i.e., IRF-to-IRF discharge) are excluded, because the IRF episode of care had not ended. We also excluded certain discharge status codes on the IRF FFS claim that indicated that the patient was not appropriate for community discharge (e.g., discharges against medical advice). See section S.9 for detailed rationale and data sources for each exclusion.

Measure exclusion criteria are as follows:

- Age under 18 years;
- No short-term acute care hospital discharge within the thirty days preceding an IRF admission;
- Discharges to a psychiatric hospital;
- Discharges against medical advice;

- Discharges to disaster alternative care site or a federal hospital;
- Discharges to court/law enforcement;
- Discharges to hospice or patient stays with a hospice benefit in the 31-day post-discharge window;
- Planned discharges to an acute or LTCH setting;
- Stays for patients without continuous Part A FFS Medicare enrollment during the 12 months prior to the IRF admission date and the 31 days after the IRF discharge;
- IRF stays preceded by a short-term acute care stay for non-surgical treatment of cancer;
- Stays ending in transfer to an IRF;
- Stays with problematic claims data (e.g. anomalous records for stays that overlap wholly or in part or are otherwise erroneous or contradictory; claims not paid);
- Exhaustion of Medicare Part A benefit during the IRF stay; and
- IRF stays in facilities outside of the United States, Puerto Rico, or another U.S. territory.

De.1. Measure Type: Outcome

S.17. Data Source: Assessment Data, Claims, Management Data

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?

Preliminary Analysis: New Measure

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a health outcome measure include providing empirical data that demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service; if these data not available, data demonstrating wide variation in performance, assuming the data are from a robust number of providers and results are not subject to systematic bias. For measures derived from patient report, evidence also should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

Evidence Summary

- The developer provided a logic model linking key inpatient rehabilitation facilities (IRF) sturctures and processes to the likelihood of a patient's successful discharge to the community.
- The developer reviewed additional literature supporting IRF structures and processes associated with the discharge to community outcome. The developer states that the effectiveness of these interventions suggests that improvement in discharge to community rates among IRF patients is possible through modifying provider-led processes and interventions. The developer summarized studies of specific interventions and their impacts as documented in the literature:

- High intensity therapy was associated with desirable discharge outcomes and may shorten PAC length of stay.
- Increased therapy intensity was associated with a larger proportion of patients being discharged to home.
- A strategy for whole-body rehabilitation-consisting of interruption of sedation, protocol-driven spontaneous breathing trials, and physical and occupational therapy in the earliest days of critical illness was safe and well-tolerated, and resulted in better functional outcomes at hospital discharge, a shorter duration of delirium, and more ventilator-free days compared with standard care.
- Longer IRF length of stay (LOS) based on patient severity and needs was positively associated with functional gains and likelihood of discharge to the community among severely impaired patients; however it was negatively associated with the likelihood of discharge to community for mildly or moderately impaired patietns.
- Greater functional status is associated with significantly higher rates of discharge to home. Interventions aimed at improving functional status can help improve discharge to community outcomes.
- Siebens Domain Management Model (SDMM), which focused on effective interdisciplinary communication and collaboration providing a standard format for weekly interdisciplinary team conferences, is associated with significantly higher discharges to community following intervention.
- Together three interventions showed significant increase in discharge to home, a decrease in re-hospitalizations, and a decrease in discharge to long-term care: the Standardized physician admission procedures with a goals-of-care discussion; palliative care consultation for patients with three or more hospital admissions over the prior 6 months; and bimonthly multidisciplinary root-cause analysis conferences for re-hospitalized patients.

Question for the Committee:

- *Is there at least one thing that the provider can do to achieve a change in the measure results?*

Guidance from the Evidence Algorithm

(Box 1) Yes → (Box 2) Yes→ Pass

Preliminary rating for evidence: ☒ **Pass** ☐ **No Pass**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- Developer analyzed all eligible IRF patient stays discharged from FY 2016-2017 Medicare FFS inpatient claims paid under Medicare's IRF Prospective Payment System (PPS) and included in the IRF Quality Reporting Program (QPP). All IRFs were included, provided they had eligible stays. Facility-level number of IRF stays ranged from 3 to 6,010 with a mean of 506.0 and standard deviation (SD) of 495.4.
- Observed scores ranged from 29.90% to 100.00%, with a mean of 65.28% and SD of 7.60 percentage points. Risk-standardized performance scores maintained a wide range, from 43.53% to 83.35%, with a mean of 64.74% and SD of 5.36 percentage points.
- Observed DTC scores by decile were:

	Min	10th pct	20th pct	30th pct	40th pct	50th pct	60th pct	70th pct	80th pct	90th pct	max
Observed DTC	29.9%	55.9%	59.6%	61.9%	64.0%	65.7%	67.3%	69.1%	69.2%	71.2%	100%
Risk-Standardized DTC	43.5%	58.0%	60.5%	62.2%	63.7%	65.2%	66.4%	67.7%	69.2%	71.1%	83.4%

- The developer compared the 95% confidence interval of each provider's performance score to the national stay-level observed DTC rate (64.82%) to determine if the provider's performance was significantly different from the national rate. Overall, 95.1% (n = 1,088) of IRFs had performance scores that were significantly different from the national rate, with 44.9% (n = 514) being worse and 50.2% (n = 574) being better than the national rate.

Disparities

- The developer assessed disparities in performance for the following social risk factors: dual eligibility, race/ethnicity, urbanicity based on beneficiary residence, and socioeconomic status (SES) using the Agency of Healthcare Research and Quality's SES Index. Some variation in patient-level DTC rates was seen across social risk factor subgroups, with the largest differences seen based on dual status.
    - The successful DTC rate was 59.5% for duals with full Medicaid, 63.1% for duals without full Medicaid, and 65.7 % for non-duals.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?

Preliminary rating for opportunity for improvement: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

Evidence: *1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure."*

- Evidence is sufficient
- Yes, the developer provided both a logic model and additional literature supporting key processes and structures within IRFs that would lend themselves to the likelihood of a successful discharge to the community.
- No additional comments
- Evidence are mostly tangential, in that the strongest evidence comes from SNF settings, not specifically IRF. And that strong SNF data specifically focuses on therapy intensity, not other processes or structures. The IRF-specific evidence was quite weak (single site; poor study design). I am skeptical about the evidence from acute care hospital discharges applying in IRF setting. That being said, if evidence is, as developers assert, likely to apply across PAC settings, then there is at least strong evidence that therapy intensity is linked to successful discharge to the community, and therapy intensity certainly is in control of facility.

Performance Gap: *(1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?)*

- High opportunity for improvement

- Yes, the developer provided an analysis of eligible IRF discharges to community and there is significant variation in performance. The developer also assessed disparities for a number of social risk factors - finding the greatest differences were seen in dual eligible populations.
- Unsure about the signifant difference noted between performance scores and the national rate?
- Yes, a performance gap is clearly shown by the wide variation among IRFs in data used for measure testing and public reporting. Data on population subgroups suggests that some disparities in DTC measure may exist, primarily related to dual-eligibility status and potentially related to race/ethnicity, although the latter relationship appears complex and likely warrants further investigation.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications **and** Testing

2b. Validity: Testing**;** Exclusions**;** Risk-Adjustment**;** Meaningful Differences**;** Comparability**;** Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

2a2. Reliability testing demonstrates if 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 enough to distinguish differences in performance across providers.

Validity

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

2b2-2b6. Potential threats to validity should be assessed/addressed.

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

Complex measure evaluated by Scientific Methods Panel? ☒ **Yes** ☐ **No**

Evaluators: Larry Glance, Karen Joynt Maddox, Marybeth Farquhar, Christie Teigland

Evaluation of Reliability and Validity:

Reliability

- Data element:
  - Did not meet NQF's requirements for data element testing. However, measure score testing was provided so this measure was still determined to be reliable.
    - The developer states that data element reliability is inferred based on the assumption of inherent accuracy of Medicare claims which are used for reimbursement and the use of claims in other NQF-endorsed measures. However, this inference does not suffice to meet NQF requirements for data element validity unless the methods and results from submission materials of these other measures are included and a summary of the analysis is provided.
    - The developer also stated auditing programs are used to assess accuracy of claims data, however, did not provide the results. While auditing is an appropriate technique

to assess data element accuracy, this does not meet NQF requirements for data element testing if results of the audit are not provided.

- Score-level:
  - Score-level reliability was conducted using split sample ICC and signal to noise approach; results indicated *"excellent"* ICC's between samples and "*good to excellent*" mean signal to noise ratios.
  - Method of signal-to-noise analysis is different from what NQF usually receives (i.e., rate divided by width of 95% CI of rate) but still acceptable.

Validity

- Data element:
  - Data element testing was done comparing two gold standard authoritative data sources for discharge to community setting using IRF-PAI (Patient Assessment Instrument) and IRF claims data which showed a high rate of agreement.
- Score-level:
  - Measure score validity was demonstrated by testing whether a facility's performance and percentile rank on the successful discharge to community measure was correlated with its performance and percentile rank on three claims based measures.
  - Face validity was only conducted on the measure concept and results of systematic collection of input was not provided.
  - Concern from many that the risk adjustment model did not include dual status although it was statistically significant in the model

Standing Committee Action Item: Discuss if dual status should be included in risk adjustment model.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The [staff] or [Scientific Methods Panel] is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The Scientific Methods Panel is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

Preliminary rating for reliability:	☒ **High**	☐ **Moderate**	☐ **Low**	☐ **Insufficient**
Preliminary rating for validity:	☐ **High**	☒ **Moderate**	☐ **Low**	☐ **Insufficient**

Evaluation A: Scientific Acceptability

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 3479

Measure Title: Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRF)

Type of measure:

☐ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☒ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☒ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☒☐ **Management Data**

☒☐ **Assessment Data** ☐ **Paper Medical Records** ☐☐ **Instrument-Based Data** ☐ **Registry Data**

☐☒ **Enrollment Data** ☒☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**

☐ **Population: Community, County or City** ☐ **Population: Regional and State**

☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐☒ **New** ☐☐ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☐☒ **Yes** ☒☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

Note: only claims information is used for measure specification. IRF-PAI data are used for "measure testing" per their document.

2. **Briefly summarize any concerns about the measure specifications.**

Numerator definition uses the terminology: "risk-adjusted predicted estimate of the number of patients", and then describes a methodology based on the observed number of actual discharges is adjusted by removing patients with unplanned admissions or death w/in 31 days. The remaining value is then "risk-adjusted for patient characteristics and a statistical estimate of the facility effect beyond case-mix (described below)." There appear to be many moving parts in how the numerator is created and highly dependent on the quality of the risk adjustment and predictionof the estimated value.

The denominator definition has similar problems including a long list of exclusions (with rationales provided) that may limit reportability. Additionally, the use of "two years of data" to ensure an "adequate number" needs more concrete specification of "adequate number" (e.g., >20? >30? >100?).

Given that the source data are coming from claims data and the "risk-adjusted predicted estimates" are being computed using a computer-based algorithm, the numerical values are "reliable" in the sense that they are repeatable.

My concern is perhaps more appropriately related to the validity of comparing two computed values (estimates) both of which are dependent upon predictive models to represent a simple concept (discharged (successfully) to the community—yes or no).

While the developer refers to Medicare FFS beneficiaries, they should include an age range for clarity.

None

none

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐☒ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure**
   ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?

   ☐ **Yes** ☐ **No**
6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

There was no obvious reliability testing of the computational methodologies used to compute the numerator or denominator or the overall measure value. A simple, randomized split-half computation at the population level would be sufficient.

The data elements are from Medicare FFS claims whose reliability has been shown and they are subject to audit. The developers conducted two types of performance measure score reliability testing: (i) measure score repeatability by assessing agreement between a facility's performance measure scores based on randomly-split independent patient stay subsets; and (ii) signal-to-noise ratios. They also stratified the facilities into quartiles based on sample size and calculated signal-to-noise ratios within each quartile. The purpose of this analysis was to assess whether measure precision was acceptable across all providers, irrespective of sample size.
Testing approaches were sufficient.

- Used only IRFs with a minimum of 25 patient stays in a two-year period. Used auditing for the reliability testing of critical data elements. Used signal-to-noise ratio to determine measure score precision. Also used split-sample reliability testing to determine measure reliability. Methods are appropriate.
- Data elements
  - Did not conduct empiric reliability testing but provided supporting evidence that claims data is reliable by stating that claims data is used in other NQF-endorsed measures and has been shown to be reliable
- Score level:
  - split-sample approach
  - signal-to-noise ratio at facility level – used a non-standard approach in which they defined SNR as the ratio of risk standardized rate to the difference between upper and lower bounds of 95% CI
- KJM: Data element testing
- Signal to noise
- Split-sample ICC

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

The mean signal-to-noise ratio in the overall sample was 77.0 with a range of 23.9 to 92.0. The mean signal-to-noise ratio was lowest in the first sample size quartile and increased progressively with increasing quartile.

The performance measure score ranged from 43.53 to 83.35 percentage points. The ICC in the overall sample was 0.77, with a 95% confidence interval (CI) of 0.74 to 0.79. The ICC was lowest in the first sample size quartile and increased progressively with increasing quartile. Overall, reliability testing results indicated good to excellent performance measure score reliability.

The results from the risk adjustment model were presented and the methodology for applying the risk-adjustment modeling (S.14) were presented. However, neither addresses the requirement to measure the reliability of computing the numerator, denominator, nor overall measure.

Regarding the risk adjustment model, the narrative uses the term "length of stay" but the EXCEL spreadsheet indicates the use of the prior length of stay in a non-psychiatric hospital as well as any stay in a psychiatric hospital as a risk factor. Although the large majority of risk factors have a $p < 0.0001$, there are several that do not. The author is not clear if all listed risk factors are used in the model (i.e., both those that are statistically significant as well as those that are not) or if only the statistically significant ones are used. Finally, narrative's reference to a "facility effect" needs to be more precisely written to know if this is a risk factor variable or an element in the logit computation (see Step 4, equation 2), or the result of applying the risk adjustment (predicted value) to the observed rate to identify differences among facilities.

- Data elements
    - Did not conduct reliability testing, but provided supporting evidence that claims data is reliable by stating that claims data is used in other NQF-endorsed measures. This is acceptable.
- Score level:
    - split-sample approach is substantial evidence score reliability
    - signal-to-noise ratio at facility level – used a non-standard approach in which they defined SNR as the ratio of risk standardized rate to the difference between upper and lower bounds of 95% CI. Difficult to interpret non-standard approach to SNR calculation. But split-sample approach is enough to provide good evidence of score reliability.

The analysis results is accurate and appropriate.

Data elements well-matched.

Signal to noise median 78 overall, by quartile of volume 74-78-80-81

ICC overall 0.77, by quartile of volume 0.62, 0.73, 0.77, 0.89

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

    Submission document: Testing attachment, section 2a2.2

    ☒☐**Yes**

    ☒☐**No**

    ☐**Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

    Submission document: Testing attachment, section 2a2.2

    ☐☒**Yes**

    ☒☐**No**

    ☐**Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and <u>all</u> testing results):

    ☐☒**High** (NOTE: Can be HIGH <u>only if</u> score-level testing has been conducted)

    ☒☐**Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has <u>not</u> been conducted)

    ☐**Low** (NOTE: Should rate <u>LOW</u> if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

☒☐**Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    Reliability testing results indicated good to excellent performance measure score reliability. The mean signal-to-noise ratio was very strong in the overall sample, with the signal being 77 times as strong as the noise, on average. The lowest mean signal-to-noise ratio was 70.9 in the lowest sample size quartile, indicating excellent signal strength even in facilities with the smallest sample sizes. The confidence interval of the performance measure score had a very narrow range, particularly in comparison with the range of the measure scores. These findings indicate that the performance measure score has excellent precision and largely represents systematic differences in quality of care across the measured entities rather than random variation or error.

    While the narrative contained a great deal of information about how the rather complex (overly complex?) numerator (predicted), denominator (expected), measure (ratio of predicted to expected), and risk adjustment model used, there was no supporting information related to the reliability of these components over time or between equivalent groups that would have provided some evidence of reliability.

    Overall the mean signal-to-noise ratio was 77.0 (increased progressively with increasing quartile numbers). The split-half reliability testing yielded an ICC of 0.77 in the overall sample indicating "good" correlation.

    Moderately reliable (and highly reliable for larger facilities)

    ICC in split-sample testing was 0.77 which suggests substantial measure reliability based on Landis scale.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.

- I am concerned by the large number of exclusions; I do not believe I have the clinical background to comment on some of these exclusions. The large number of exclusions appears to lead to the need for two years of data for each institution before there are sufficient numbers to perform the computations described. My concern is that the results of this two-year lag does not provide timely information to the health care provider regarding any changes in practice that were implemented to improve patient discharge rates.
- No concerns.

None

The exclusions are well rationalized and documented with literature and appear to be reasonable.

- 14% of cases were excluded due to no prior inpatient hospitalization. Need data from prior hospitalization for comorbidity information. Could have considered using carrier data to get this comorbidity data.
- Similar measures from this group also had ~10% of cases excluded because of lack of continuous enrollment in Medicare. This is a limitation of this measure, and the other measures in this "group" (3477)

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.

    I did not find any information provided by the author that described meaningful differences among these health care providers.

No concerns.

None.

95% facilities were identified as quality outliers, suggesting that measure can identify meaningful differences in performance.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
Submission document: Testing attachment, section 2b5.

Na

None.

While several data sources were identified by the author for the purpose of "testing" the measure, results of this testing were not included in the materials.

n/a

15. **Please describe any concerns you have regarding missing data.**

Submission document: Testing attachment, section 2b6.

Claims data are usually quite complete; hence, the source data for the measure are not likely to be missing.

None.

Rare to have missing data. Developer cited missing data were only present in 3 patient stays, and therefore, does not have an impact on the measure.

none

No concerns. Missing data fraction is negligibly small.

16. **Risk Adjustment**

16a. **Risk-adjustment method** ☐ **None** ☒ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒☐ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☒ No ☐ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐☒ Yes ☒☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐☒ Yes ☒☐ No

Risk standardized rates were compared overall with only <1% difference in RSR for entire sample when comparing RSR with and without SES adjustment. Should have considered comparing RSR with and without SES adjustment for individual facilities using ICC analysis. Agree that effect of SES adjustment is small.

16d.**Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☒ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☒☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐☒ Yes ☒☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☒ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐☒ Yes ☒☐ No

16e**. Assess the risk-adjustment approach**

The testing for the impact of the ONE SES variable tested, dual eligible status, indicated a disparity in rates for duals and non-duals, rates of 75.2% vs. 71.3% AFTER adjusting for all other risk factors. The coefficient on the full dual indicator was -0.2098 with small SE (0.0095) and p-value <.0001 and OR of 0.811, indicating a nearly 20% disparity in rates.

- The rationale provided for not including dual status in the model was the low impact of dual status in the risk adjustment model. I am not surprised that the impact on the overall mean rates and model performance did not change significantly with addition of one variable given there are 288 variables in the model. I DO NOT think this indicates that the additional adjustment for **Dual status** could not potentially generate more accurate and comparable rates for both plans that serve a LARGE portion of duals (their rates may look worse than they actually are) and those having a SMALL proportion of duals (their rates may appear better than they actually are based on their wealthier healthier population). They further claim adjusting for dual status may "mask potential disparities in quality" even though a hierarchical model with significant risk adjusters for clinical risk factors was utilized and should control for between facility differences which may be attributable to quality per CMS and RAND previous guidance.
- Although the c-statistic provided in the risk adjustment table is adequate >0.7, there are concerns about the specificity of about which risk factors are or are not included in the model. Additionally, there is a comment in the narrative that the risk model is recomputed each quarter. Why? How do the models differ over time (e.g., significance of risk factors, changes in risk factors, predictive power)?
- Risk-adjustment approach is similar to other discharge to community measures and is appropriate.
- Overall this is an excellent risk adjustment model, and very thoughtfully developed. I disagree with the handling of dual status – there is a meaningful relationship with the outcome and thus it seems to me that dual status should be included in the model. Whether or not a lot of facilities change scores, or whether the c-statistic changes in the model, is not the appropriate test for impact; instead, the impact on the tails should be examined.
- Hierarchical GLM
- Created non-parsimonious risk adj model, which is appropriate for a risk adj model to ensure adequate case mix adjustment, and includes principal diagnoses from the prior short-term acute care stay; IRF case-mix groups (CMG); length of stay, types of surgery or procedures (AHRQ CCS), and dialysis from the prior short-term acute care stay; comorbidities (CMS HCC); and number of prior hospitalizations in the year preceding the IRF admissio
- Quality is quantified using PE ratio
- Model includes following variables from preceding inpatient stay
- Model performance was not reported for validation data set, and appears to have been done using the same data used to develop model
    - C statistic of 0.71 is acceptable and is slightly better than many readmission models which tend to have low C stats (note, a C stat of 0.5 means model is no better than the flip of a coin).
    - Calibration – as assessed using calibration curves/OE ratios by deciles is excellent

For cost/resource use measures ONLY:

17. **Are the specifications in alignment with the stated measure intent?**

    ☐ **Yes** ☐ **Somewhat** ☐ **No (If "Somewhat" or "No", please explain)**

18. **Describe any concerns of threats to validity related to attribution, the costing approach, carve outs, or truncation (approach to outliers):**

VALIDITY: TESTING

19. **Validity testing level:** ☐☒ **Measure score** ☐☒ **Data element** ☒☐ **Both**

20. **Method of establishing validity of the measure score:**

    ☐☒ **Face validity**

    ☐☒ **Empirical validity testing of the measure score**

    ☒☐ **N/A (score-level testing not conducted)**

21. **Assess the method(s) for establishing validity**

    Submission document: Testing attachment, section 2b2.2

Extensive testing was completed to validate the adequacy of the risk adjustment model.

My concern is related to the validity. The measure is a comparison of two computed values (numerator = predicted; denominator = expected) with both values using a predictive model and then applying different exclusions. How is the ratio of two predicted values better than a simple observed discharged rate (successful = numerator; total = denominator) to the community with the health care provider's overall rate adjusted for the patients' characteristics that they serve?

- Used a variety of methods to establish validity, all are appropriate.
- Validity of data elements assessed by comparing key data elements in database to authoritative source – this was done for outcome variable
- Face validity score assessed using TEP
- Assessed convergent validity by comparing to other measures
- Empiric validity is assessed by assessing predictive validity of risk adjustment model

Both Face validity and empirical validity (convergent, known-groups) very well done

22. **Assess the results(s) for establishing validity**

    Submission document: Testing attachment, section 2b2.3

- All testing led similar conclusions and indicated strong validity of the measure
- The narrative contained references to literature that were related to validity. There was no empirical evidence of the how validity was established or tested.
- Test results indicate that both data element and measure score are valid.
- Highly valid
- Results from TEP indicate that measure concept is important – but no documentation that TEP evaluated face validity of measure & approach to risk adjustment, other than the concept of measuring this outcome
- Demonstrated accuracy of outcome variable compared to authoritative source & indicated excellent agreement (99.2%)
- Convergent validity – showed weak inverse correlation with (1) potentially preventable readmissions, (2) potentially preventable within-stays readmissions; and (3) Medicare spending per bene
- Empiric testing of risk adjustment model indicates very good model performance

23. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

    Submission document: Testing attachment, section 2b1.

    ☐☒**Yes**

    ☐**No**

    ☒☐**Not applicable** (score-level testing was not performed)

24. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?** *NOTE that data element validation from the literature is acceptable.*

    Submission document*: Testing attachment, section 2b1.*

    ☐☒**Yes**

    ☐**No**

    ☐**Not applicable** (data element testing was not performed)

25. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

    ☐☒**High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒☐**Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

    ☒☐**Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

    ☐**Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

26. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

    See 16e above.

    The supporting documentation reviewed did not provide adequate evidence empirical

    Through testing of data elements and measure score for validity.

    I disagree with the handing of dual status but otherwise think this is a very well-developed measure

    Performance of risk adjustment model – which is a predictive validity of measure – is very good

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

27. **What is the level of certainty or confidence that the empirical analysis demonstrates that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct?**

    ☐**High**

    ☐**Moderate**

    ☐**Low**

    ☐**Insufficient**

28. **Briefly explain rationale for rating of EMPIRICAL ANALYSES TO SUPPORT COMPOSITE CONSTRUCTION**

ADDITIONAL RECOMMENDATIONS

29. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**

Dual status

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

Reliability- Specifications: *(2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm*

or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?)

- Data elements reliable
- I have concerns regarding the developer's assumption of accuracy in Medicare claims data to be used exclusively for this measure, given that their purpose is largely for reimbursement. Additionally, as the NQF staff review indicates, this inference does not meet NQF data requirements for data element testing. However, this does still seem a reasonable approach - therefore, I would rate it as moderate for reliability (particularly given my concerns about exclusive reliance on claims data). Another important point is the many components to how the numerator is constructed and also the large number of exceptions in the denominator.
- Should the committee have concerns about not meeting the NQF's requirements for data element testing?
- Specifications are clear

Reliability-Testing: *"2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?"*

- Testing reliable
- I may have addressed some of this above, but I would also like to agree with the NQF assessment that dual status should be included in the risk adjustment for this measure.
- Yes? As noted previosly, should the committee be concerned about the NQF's requirements for data element testing?
- No concerns

Validity-Testing: *"2b1. Validity -Testing: Do you have any concerns with the testing results?"*

- Maybe. Dual eligible issue though model is only slightly impacted
- As noted above, risk adjustment for dual eligibles should be included. Additionally, there are a large number of exclusions; however, they do seem to be appropriate ones. Agree with the NQF assessment of moderate.
- No concerns
- No concerns

Validity- Threats to Validity: *"2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?"*

- Need to come to consensus to risk adjust for dual eligibles. Current position is to exclude.
- No major concerns here other than the need to risk adjust for dual eligibles.
- No comments
- I share some of the methods panelists' questions about the large number of measure exclusions. I would like to see some data on the % of cases excluded by each exclusion criterion. I also agree that requiring 2 years of data undermines the timeliness and actionability of the measure, so if the requirement for 2 years of data stems from the large number of exclusions, then that is a real element of concern.

Other Threats to Validity: *"2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment)2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure?2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?"*

- Excusions appropriate and risk adjustmnt model acceptable
- As noted earlier, there are a large number of exclusions - but they seem appropriate. The measure should be risk adjusted for dual eligibles.

- No comments
- I agree our committee should discuss potential risk adjustment for dual eligible status.

Criterion 3. Feasibility

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- This measure uses Medicare Part A inpatient claims data, which are routinely collected for payment purposes. These data are electronically available from the Centers for Medicare & Medicaid Services (CMS) at no cost beyond that of data processing and can be used to specify, publicly report, and track the measure in a timely fashion.
- All data elements are in defined fields in electronic claims and generated or collected by and used by healthcare personnel during the provision of care. The data are coded by someone other than person obtaining original information.
- There are no fees, licensing, or other requirements to use any aspect of the measure as specified.

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?
- Is the data collection strategy ready to be put into operational use?

Preliminary rating for feasibility: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

Feasibility: *"3. Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?"*

- Feasible
- Feasibility seems very high given that it comes from a single data source that is already being collected as part of the course of care (for reimbursement).
- This is feasible as all data is electronically available
- Data are all readily available. My only feasibility concern is around requirement for 2 years of data; this will make the measure always out-of-date by the time it's calculated. What's the value in that?

Criterion 4: Usability and Use

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

Accountability program details

- Inpatient Rehabilitation Facilities Quality Reporting Program

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- Confidential feedback reports on the DTC-PAC IRF measure were provided to all active IRF providers under the IRF PPS in October 2017. Active providers received provider preview reports in June 2018, prior to public reporting of the measure in September 2018.
    - Confidential feedback reports included the following data: provider number, DTC reporting period start date, DTC reporting period end date, observed number of discharges to community, number of eligible stays, observed discharge to community rate, risk-standardized discharge to community rate, national observed discharge to community rate, comparative performance category, number of IRFs that performed better than the national rate, number of IRFs that performed no different than the national rate, number of IRFs that performed worse than the national rate, and number of IRFs too small to report.
- The developer also solicited public comments on this measure via a 30-day public comment period during November-December 2015, and during the FY 2017 IRF PPS rulemaking process.
    - The developer cited the Post Acute Care Quality Initiatives Public Comments Summary Report for the Development of a Discharge to Community Measure report for documentation of feedback on the measure. The developer did not provide a summary of the feedback but noted that the measure received extensive support.

Questions for the Committee:

- How has the measure been vetted in real-world settings by those being measured or others?

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- This measure was implemented in October 2016 and was publicly reported for the first time in September 2018 using FY 2016-2017 data. Therefore, the developer noted that there is no data to assess trends in performance over time.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- The developer reports no unexpected findings during implementation of the measure.

Potential harms

- The developer reports that no unintended impacts on patients have been detected to date.

***Questions for the Committee**:*

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- Do the benefits of the measure outweigh any potential unintended consequences?

Preliminary rating for Usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

Use: *"4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided?4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?"*

- Use PASS
- Feedback reports from those that have been using the measure were provided and public comments were solicited.
- Feebdack from the public was obtained.
- IRFs had opportunities to see their measure results in advance of public reporting, but it's not clear the extent to which any feedback provided was acted on. Similarly, the report on public comments contains many reasonable concerns raised by stakeholders, yet in a preliminary skim, I didn't see any cases where the measure developers acted on this feedback. Rather, the report only offers rationale for NOT making changes. This leaves much potential room for improvement in the measure.

Usability: *"4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations?4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them."*

- Usability is moderate
- I do not think it is clear yet if these performance results can be used to truly further the goal of high-quality, efficient care. The concept behind the measure is good and it seems relatively well constructed, but also seems overly complex for what it is trying to measure. This also makes it less clear to me if there are any potential unintended consequences as of yet. Would rate this as moderate.
- Data could be used to evaluated IRF discharges which may be beneficial to various stakeholders
- The measure is already in use for public reporting but it's unclear what plans are in place for use in quality improvement, nor how QI efforts would be meaningful when based on measure results requiring 2 years of data, and therefore quite out-of-date.

Criterion 5: Related and Competing Measures

Related or competing measures

- The developer did not identify any measures that are related or competing to this measure.
- This measure is related to, but not competing with, three measures that were submitted by the same developer for review in this cycle:
    - 3477: Discharge to Community-Post Acute Care Measure for Home Health Agencies
    - 3480: Discharge to Community-Post Acute Care Measure for Long-Term Care Hospitals (LTCH)
    - 3481: Discharge to Community-Post Acute Care Measure for Skilled Nursing Facilities (SNF)

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

Related and Competing: *"5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?"*

- No
- No additional measures to identify here
- None noted
- no concerns

Public and Member Comments

NQF received no comments on this measure as of: February 1, 2019

Brief Measure Information

NQF #: 3479

Corresponding Measures:

De.2. Measure Title: Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRF)

Co.1.1. Measure Steward: Centers for Medicare & Medicaid Services

De.3. Brief Description of Measure: The Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (DTC-PAC IRF) was developed to address the resource use and other measures domain of Discharge to the Community mandated by the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This outcome measure assesses successful discharge to community from an IRF, with successful discharge to community including no unplanned rehospitalizations and no death in the 31 days following IRF discharge. The measure reports an IRF's risk-standardized rate of Medicare fee-for-service (FFS) patients who are discharged to the community following an IRF stay, and do not have an unplanned readmission to an acute care hospital or long-term care hospital (LTCH) in the 31 days following discharge to community, and who remain alive during the 31 days following discharge to community. The measure is calculated using two consecutive years of Medicare FFS claims data and was developed using calendar year (CY) 2012-2013 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., IRF discharges from October 1, 2015 through September 30, 2017.

The measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the IRF Quality Reporting Program (QRP) finalized in the FY 2017 IRF Prospective Payment System (PPS) Final Rule and implementation began October 1, 2016 [1]. Confidential feedback reports on measure performance were distributed to IRF providers in Fall 2017. The measure will be publicly reported on the IRF Compare website (https://www.medicare.gov/inpatientrehabilitationfacilitycompare/) in Fall 2018 using FY 2016-2017 data. Four claims-based discharge to community measures were developed for IRF, LTCH, skilled nursing facility, and home health agency settings to meet the mandate of the IMPACT Act. These measures were conceptualized uniformly across the four settings, in terms of the definition of the discharge to community outcome, the approach to risk-adjustment, and the measure calculation.

References

[1] Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 Federal Register, Vol. 81, No. 151. https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf

1b.1. Developer Rationale: Successful discharge to community from IRF is widely recognized as an important outcome for patients and their families and an indicator of IRF quality of care.[1-6] The Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act of 2014) mandated the specification of a measure to address the domain of discharge to community across post-acute care (PAC) settings.[4] Discharge to community is an important goal for both patients with potential for functional improvement, and patients who may not be expected to make functional gains given their clinical condition or disease. IRFs frequently track their discharge to community outcomes for internal quality improvement purposes; however, this tracking commonly assesses discharge destination alone and not post-discharge readmissions or death. By assessing whether patients successfully remain in the community without acute complications for 31 days following discharge, the DTC-PAC IRF measure is a more meaningful, patient- and family-centered measure of successful community discharge. Discharge to community is an actionable health care outcome as targeted interventions have been shown to successfully increase discharge to community rates in a variety of post-acute settings, including IRFs [7,8] (see attached evidence form). The effectiveness of these interventions suggests that improvement in discharge to community rates among IRF patients is possible through modifying provider-led

processes and interventions. In addition to being important from a patient and family perspective, avoiding institutionalization and promoting discharge to community when appropriate are also important from a cost and resource use perspective. Patients discharged to community settings, on average, incur lower costs over the recovery episode, compared with patients discharged to institutional settings.[9-11]

Measuring and publicly reporting the DTC-PAC IRF measure is expected to help differentiate IRFs based on quality of care, and drive improvement in this outcome. The DTC-PAC IRF measure would impact a large number of Medicare beneficiaries and the overall Medicare program. In 2016, Medicare spent $7.7 billion on fee-for service (FFS) IRF care; this care was provided in about 1,200 IRFs nationwide, covering more than 391,000 IRF stays for about 350,000 beneficiaries.[5] There are no NQF-endorsed measures assessing successful discharge to community from IRFs.

[1] Gillsjö C, Schwartz-Barcott D, von Post I. Home: The place the older adult can not imagine living without. BMC Geriatrics. 2011;11(1):10.

[2] Unsworth C. Clients´ perceptions of discharge housing decisions after stroke rehabilitation. American Journal of Occupational Therapy. 1996;50(3):207-216.

[3] Roush CV, Cox JE. The meaning of home: how it shapes the practice of home and hospice care. Home Healthcare Now. 2000;18(6):388-394.

[4] Improving Medicare Post-Acute Care Transformation Act of 2014. Public Law 113–185—October 6, 2014. Available at: https://www.gpo.gov/fdsys/pkg/PLAW-113publ185/pdf/PLAW-113publ185.pdf.

[5] Medicare Payment Advisory Commission. Report to the Congress: Medicare Payment Policy. March 2018. Available at: http://www.medpac.gov/docs/default-source/reports/mar18_medpac_entirereport_sec.pdf.

[6] Hsieh C-H, DeJong G, Groah S, Ballard PH, Horn SD, Tian W. Comparing rehabilitation services and outcomes between older and younger people with spinal cord injury. Archives of Physical Medicine and Rehabilitation. 2013;94(4, Supplement):S175-S186.

[7] Kushner DS, Peters KM, Johnson-Greene D. Evaluating Siebens Domain Management Model for inpatient rehabilitation to increase functional independence and discharge rate to home in geriatric patients. Archives of Physical Medicine and Rehabilitation. 2015;96(7):1310-1318.

[8] Kushner DS, Peters KM, Johnson-Greene D. Evaluating use of the Siebens Domain Management Model during inpatient rehabilitation to increase functional independence and discharge rate to home in stroke patients. PM&R: the Journal of Injury, Function, and Rehabilitation. 2015;7(4):354-364.

[9] Dobrez D, Heinemann AW, Deutsch A, Manheim L, Mallinson T. Impact of Medicare´s prospective payment system for inpatient rehabilitation facilities on stroke patient outcomes. American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists. 2010;89(3):198-204.

[10] Gage B, Morley M, Spain P, Ingber M. Examining Post Acute Care Relationships in an Integrated Hospital System Final Report. RTI International; 2009.

[11] Medicare Payment Advisory Commission. Paying for sequential stays in a unified prospective payment system for post-acute care. In: Report to the Congress: Medicare and the Health Care Delivery System. June 2018. Available at: http://www.medpac.gov/docs/default-source/reports/jun18_ch4_medpacreport_sec.pdf?sfvrsn=0

S.4. Numerator Statement: The measure numerator is the risk-adjusted predicted estimate of the number of patients who are discharged to the community, and do not have an unplanned readmission to an acute care hospital or LTCH in the 31-day post-discharge observation window, and who remain alive during the post-discharge observation window.

This estimate starts with the observed number of discharges to community, defined as:

(i) discharges to home or self care with or without home health services, based on Patient Discharge Status Codes 01, 06, 81, or 86 on the Medicare FFS claim [2]; and

(ii) no unplanned acute or LTCH hospitalizations in the 31-day post-discharge window; and

(iii) no death in the 31-day post-discharge window.

The discharge to community outcome is risk-adjusted for patient characteristics and a statistical estimate of the facility effect beyond case-mix (described below).

References

[2] National Uniform Billing Committee Official UB-04 Data Specifications Manual 2018, Version 12, July 2017, Copyright 2017, American Hospital Association.

S.6. Denominator Statement: The target population for the measure is the group of Medicare FFS beneficiaries who are discharged from an IRF during the measurement period and are not excluded based on the measure exclusion criteria (see S.8. and S.9.).

The measure denominator is the risk-adjusted expected number of discharges to community. This estimate includes risk-adjustment for patient characteristics with the facility effect removed. The "expected" number of discharges to community is the predicted number of risk-adjusted discharges to community if the same patients were treated at the average facility. The logistic regression model used to calculate the denominator is developed using all non-excluded facility stays in the national data. The denominator is computed in the same way as the numerator, but the facility effect is set at the average.

S.8. Denominator Exclusions: Measure exclusion criteria are based on administrative data from Medicare claims and eligibility files. Exclusion criteria were selected to maintain clinical validity of the measure by excluding stays for which discharge to community would not be appropriate, to ensure data availability and completeness, to exclude stays with problematic claims data, and to maintain relevance to the IRF QRP (e.g., excluding IRFs not included in the IRF QRP based on regional location). Only IRF stays that are preceded by a short-term acute care stay in the 30 days prior to the IRF admission date are included in the measure; this is because risk-adjustment variables come from the short-term acute care stay in the 30 days prior to IRF admission. Stays ending in transfers to the same level of care (i.e., IRF-to-IRF discharge) are excluded, because the IRF episode of care had not ended. We also excluded certain discharge status codes on the IRF FFS claim that indicated that the patient was not appropriate for community discharge (e.g., discharges against medical advice). See section S.9 for detailed rationale and data sources for each exclusion.

Measure exclusion criteria are as follows:

- Age under 18 years;
- No short-term acute care hospital discharge within the thirty days preceding an IRF admission;
- Discharges to a psychiatric hospital;
- Discharges against medical advice;
- Discharges to disaster alternative care site or a federal hospital;
- Discharges to court/law enforcement;
- Discharges to hospice or patient stays with a hospice benefit in the 31-day post-discharge window;
- Planned discharges to an acute or LTCH setting;
- Stays for patients without continuous Part A FFS Medicare enrollment during the 12 months prior to the IRF admission date and the 31 days after the IRF discharge;
- IRF stays preceded by a short-term acute care stay for non-surgical treatment of cancer;
- Stays ending in transfer to an IRF;
- Stays with problematic claims data (e.g. anomalous records for stays that overlap wholly or in part or are otherwise erroneous or contradictory; claims not paid);
- Exhaustion of Medicare Part A benefit during the IRF stay; and
- IRF stays in facilities outside of the United States, Puerto Rico, or another U.S. territory.

De.1. Measure Type: Outcome

S.17. Data Source: Assessment Data, Claims, Management Data

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

NQF_DTC-PAC_IRF_Evidence_Form_RTI.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:**

Measure Title: Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRF)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here:

Date of Submission: 12/18/2018

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *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.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

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

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Outcome: 3 Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence 4 that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: 5 a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence 4 that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence 4 that the measured structure leads to a desired health outcome.
- Efficiency: 6 evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☒ Outcome: Successful discharge to community

☐Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (*A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)*

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should

be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

This claims-based outcome measure assesses successful discharge to community from an IRF setting, with successful discharge to community including no unplanned readmissions to an acute hospital or long-term care hospital (LTCH) and no death in the 31 days following discharge. Successful discharge to a community-based setting is an important goal for patients in IRF settings. Section **1a.2** below cites literature support regarding provider structures and processes associated with the discharge to community outcome. Interventions that enhance discharge transitions and coordination of care have been associated with increased discharge to community rates and reduced readmission rates following community discharge.

Figure 1 illustrates some key IRF structures and processes that influence the likelihood of a patient's successful discharge to community. The structures and processes listed in the figure are not exhaustive but are intended as examples of key structures and processes that could be implemented to facilitate successful discharge to community. **Figure 1** is based on literature examining the relationship between structures, processes and successful community transitions, as well as the Department of Health and Human Services' roadmap to better care transitions and fewer readmissions, which lists elements for safe, effective and efficient care transitions. [1]

References

1. United States Department of Health and Human Services: Roadmap to better care transitions and fewer readmissions. 2011. Available at: https://www.in.gov/isdh/files/improve_care_tran.pdf

Figure 1. Examples of IRF Structures and Processes that Influence the Outcome of Successful Discharge to Community

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

Not applicable.

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

The empirical evidence provided below comes from IRF-specific literature, as well as literature from other inpatient PAC and hospital settings, as evidence related to healthcare structures and processes for improving discharge to community outcomes is largely applicable across inpatient PAC and hospital settings. There is consistent evidence in the literature across inpatient settings that rehabilitation interventions, discharge planning, and care coordination can improve discharge to community rates. Thus, evidence from other inpatient PAC and hospital settings can be used to support the DTC-PAC IRF measure. Discharge to community is an actionable health care outcome, as targeted interventions have been shown to successfully increase discharge to community rates in a variety of post-acute settings. These interventions frequently involve specific rehabilitation strategies such as addressing discharge barriers and improving medical and functional status, discharge planning, communication and care coordination, or community-based transitional care services and supports.

In a retrospective observational study, O'Brien and Zhang [2] examined the relationship of therapy intensity with discharge destination and time to community discharge (i.e., LOS) among 311,338 Medicare fee-for-service (FFS) residents in 3,605 SNFs across five states in 2008. The authors used data from Minimum Data Set (MDS), Online Survey Certification and Reporting (OSCAR) dataset, and Rural Urban Commuting Area Codes. Therapy intensity was calculated as the total minutes of physical, occupational, and speech therapy in a day, and categorized as high (≥ 60min/day), medium-high (45 to <60 min/day), medium-low (30 to <45min/day), and low (0 to <30min/day). The authors found a dose-response relationship between therapy intensity and discharge destination, with the proportion of residents discharged to community decreasing with decreasing therapy intensity: 63% (high intensity), 52.9% (medium-high), 45.1% (medium-low), and 27.4% (low). The proportion of residents discharged to long-term care increased as therapy intensity decreased: 8.4% (high), 13.3% (medium-high), 17.1% (medium-low), and 24.4% (low). Risk-adjusted random-effects competing risks regression modeling controlling for patient demographic/clinical characteristics and facility/regional characteristics, showed that compared with the high intensity group, the medium-high, medium-low, and low intensity groups, respectively, had a 15% (hazard ratio (HR) = .85, 95% confidence interval (CI) = .83-.85), 32% (HR = .68, 95% CI = .67-.69), and 57% (HR = .43, 95% CI = .42-.45) lower likelihood of community discharge than of becoming permanently placed in a nursing home. The hazard of hospital readmission increased with decreased therapy intensity. Compared with the high intensity group, the medium-high, medium-low, and low intensity groups, respectively, had an 8% (HR = 1.08, 95% CI = 1.06-1.12), 25% (HR = 1.25, 95% CI = 1.19-1.27), and 29% (HR = 1.29, 95% CI = 1.19-1.27) higher risk for hospital discharge than for permanent nursing home placement. The risk of death also increased significantly as therapy intensity decreased (HR = 1.407, 95% CI = 1.32-1.45; HR = 2.299, 95% CI = 2.15-2.46; and HR = 4.198, 95% CI = 3.89-4.52 for medium-high, medium-low, and low intensity groups, respectively). For residents discharged home (n = 162,792), the mean SNF LOS increased as therapy intensity decreased from 35.6 ± 24.2 days for the high intensity group to 45.3 ± 31.7 days for the low intensity group. Further, Poisson regression modeling controlling for covariates and compared with the low intensity group, showed that SNF LOS was 5% shorter for the high intensity group (p<.001), with an incident rate ratio of 0.95 (95% CI = 0.92-0.97). The high intensity group averaged 2 days less in the SNF compared with other intensity groups. The authors concluded that high intensity therapy was associated with desirable discharge outcomes and may shorten PAC length of stay. [2]

In another retrospective cohort study, Jung et al. [3] examined the relationship between therapy intensity and likelihood of discharge to home in 481,908 Medicare FFS residents admitted to 15,496 SNFs after hip fracture. Therapy intensity included total physical, occupational, and speech and language therapy minutes, and was calculated as the average quantity of therapy per week. Patient-level data were taken from MDS and Medicare inpatient claims and facility-level data from OSCAR, for years 2000 through 2009. Multivariable linear regression adjusting for patient characteristics and time-varying facility characteristics indicated that each additional hour of therapy per week was associated with a 3.1 percentage-point (95% CI = 3.0, 3.1) increase in the likelihood of discharge to home. An additional hour of occupational therapy was associated with a 5.3 percentage-point (95% CI = 5.2, 5.4) increase in the likelihood of discharge to community, while an additional hour of physical therapy was associated with a 5.9 percentage-point (95% CI = 5.8, 6.1) increase in the likelihood of discharge to community. When examined by SNF LOS, an additional hour of therapy per week was associated with increases in the likelihood of discharge to home of 2.9 percentage points (95% CI = 2.8, 2.9), 3.0 percentage points (95% CI = 2.9, 3.1), and 3.0 percentage points (95% CI = 3.0, 3.1) for stays of up to 30, 60, and 90 days, respectively. The effect of additional therapy decreased as the Resource Utilization Group (RUG) category increased, with additional therapy not benefiting patients in the highest RUG category, who had the highest impairment levels. The authors concluded that increased therapy intensity was associated with a larger proportion of patients being discharged to home, suggesting better post-acute outcomes, except for patients with the highest impairment levels. [3]

Schweickert et al. [4] conducted a randomized controlled trial of physical and occupational therapy in 104 patients receiving mechanical ventilation in medical intensive care in two Midwest medical centers. Intervention group patients (n = 49) received early exercise and mobilization (physical and occupational therapy) during periods of daily interruption of sedation, while control group patients (n = 55) received daily interruption of sedation and standard care with physical and occupational therapy as ordered by the primary care team. Blinded therapists functionally assessed patients at discharge based on the ability to perform six activities of daily living and walk independently. Using intention-to-treat analysis, the authors reported higher discharge to home rates in intervention patients (43%) compared with controls (24%) for comparison of home discharge to all other possible locations for group comparison (p=0.06). Return to independent functional status at hospital discharge occurred in 59% of intervention patients compared with 35% of controls (p=0.02; odds ratio (OR) = 2.7 [95% CI = 1.2-6.1]). Other important outcomes included shorter median duration of delirium (2 vs. 4 days; p=0.02), and more median ventilator-free days (23.5 vs. 21.1; p=0.05) during the 28-day follow-up period in intervention patients than controls. The authors concluded that a strategy for whole-body rehabilitation-consisting of interruption of sedation, protocol-driven spontaneous breathing trials, and physical and occupational therapy in the earliest days of critical illness was safe and well-tolerated, and resulted in better functional outcomes at hospital discharge, a shorter duration of delirium, and more ventilator-free days compared with standard care. [4]

Length of stay (LOS) is another important variable that can impact discharge to community rates. Camicia et al. [5] examined the relationship between IRF LOS and discharge to community outcomes in a retrospective cohort analysis of 4,781 IRF patients with stroke between 2009 and 2011, based on random sampling of 2% of all stroke patients during the time period, using Uniform Data System for Medical Rehabilitation (UDSMR) data. After adjusting for admission functional status and other patient factors, IRF LOS was positively associated with functional gains and likelihood of discharge to community among severely impaired patients (OR = 1.010, 95% CI = 0.999–1.021), but negatively associated with the likelihood of discharge to community for mildly (OR = 0.905, 95% CI = 0.839–0.976) and moderately (OR = 0.943, 95% CI = 0.924–0.962) impaired patients. [5] Thus, optimizing IRF LOS based on patient severity and needs is important to improve discharge to community outcomes.

Functional status has been observed to be associated with discharge destination, including discharge to home. For example, Thrush et al. [6] examined the relationship between functional status and discharge outcomes based on data collected from 101 LTCH patients in a 38-bed LTCH over 8 months, beginning in September 2010. Functional status was measured based upon the Functional Status Score for the Intensive Care Unit (FSS-

ICU), which contains five functional activities scored using a seven-point system, resulting in a score range from 0 to 35; FSS-ICU has been used in both the ICU and LTCH setting. FSS-ICU scores were significantly higher for those discharged home (score = 28) compared to those discharged to a long-term care/hospice/expired (score = 5) or transferred to a short-stay hospital (score = 4) (p<0.001). [6] These findings suggest that interventions aimed at improving functional status could help improve discharge to community outcomes.

Using an observational study design, Kushner et al. [7,8,9] assessed the impact of the Siebens Domain Management Model (SDMM) on several discharge outcomes in IRF geriatric, stroke, and geriatric-stroke patients at a single facility, and compared outcomes to national IRF outcomes using UDSMR data. The SDMM intervention focused on effective interdisciplinary communication and collaboration providing a standard format for weekly interdisciplinary team conferences. The intervention also involved weekly adjustments of care focusing on potential barriers to home or community discharge including medical/surgical issues, mental status/emotions/coping, physical function, and living environment/community re-entry needs. In all three patient groups, the authors reported significantly higher discharges to community in the post-intervention period (year 2012) compared with pre-intervention (year 2010) (p < 0.05). Pre-intervention versus post-intervention discharge to community rates were 58.5% (of 429) vs. 74.4% (of 524) in geriatric patients [7], 57.8% (of 154) vs. 81.2% (of 151) in stroke patients [8], and 56.9% (of 60) vs. 79.3% (of 58) in geriatric-stroke patients [9]. The authors also reported other outcome improvements following SDMM implementation including fewer discharges to long-term care (24.0% pre-intervention vs. 10.4% post-intervention) [7], fewer acute care transfers (27.3% pre-intervention vs. 9.4% post-intervention) [8], reduced length of IRF stay [2,8,9], and improved Functional Independence Measure (FIM) efficiency [7,8]. While the authors did not adjust for patient characteristics when comparing outcomes, the magnitude of differences strongly suggests that discharge planning processes can improve discharge to community rates. The authors also reported that unlike the pre-intervention group, the post-intervention group had significantly higher (3-4 times higher) discharge to community rates [7,8,9], fewer acute care transfers [7,9], fewer long-term care discharges [9], and higher FIM efficiency [7,8,9] compared with case-mix group adjusted national UDSMR data, using a 0.05 significance level.

Berkowitz et al. [10] examined the impact of a three-component intervention on discharge disposition outcomes of residents admitted to a single SNF between June 2008 and May 2010. The intervention included standardized physician admission procedures with a goals-of-care discussion; palliative care consultation for patients with three or more hospital admissions over the prior 6 months; and bimonthly multidisciplinary root-cause analysis conferences for rehospitalized patients to identify problems and improve processes of care. 862 patients were included in the pre-intervention period (June 2008–May 2009) and 863 during the post-intervention period (June 2009–May 2010). Discharge dispositions differed significantly (p =.03) between the pre- and post-intervention periods, with discharges to home increasing from 68.6% to 73.0%. The rate of rehospitalization declined 19.4% from 16.5% to 13.3%, and discharges to long-term care fell from 13.8% to 11.5%. [10]

Buttke et al. [11] described outcomes of Minnesota's Return to Community Initiative (RTCI), a novel, statewide initiative introduced in 2010 to assist private paying nursing home residents to return to and remain in the community without converting to Medicaid. RTCI is a multi-component intervention, consisting of in-person SNF visits to ensure consumers receive information regarding options for residing in the community and make consumers aware of the right to live in the least restrictive environment; interviews to fill the Community Planning Tool; development of Community Living Support Plans; post-discharge in-person and phone follow-up visits to help the consumer transition back to the community; ongoing follow-up up to 90 days post-discharge; and follow-up up to five years if desired by the consumer. The authors reported that under RTCI, the number of resident transitions to the community increased from 38 per month in 2013 to 69 per month in 2014, and 90 to 100 per month during 2015 and 2016. Seventy-six percent of transitioned residents were alive and living in the community at one year after initial transition. The authors concluded that the relatively low nursing home readmission rates and mortality among transitioned residents may be attributable to effective follow-up. [11]

In a retrospective analysis, Logue and Drago [12] described the impact of a modified community-based care transitions program on 30-day all-cause readmissions in 149 Medicare FFS patients in two hospital catchment areas in Arizona. The care transitions program included home-based in-person and phone visits by licensed practical nurses and registered nurses. The program focused on medication self-management, use of a personal health record by the patient or caregiver to facilitate communication and ensure continuity of the care plan across providers and settings, timely follow-up visits with care teams, educating patients on red flags indicating worsening condition, and depression and mobility screening. The 30-day all-cause readmission rate was 4% for patients who completed the program; compared with a baseline readmission rate of 15%, the program resulted in a 73% reduction in all-cause readmissions. Compared with the national average 30-day readmission rate, the program resulted in an 80% reduction in readmissions. The authors also reported other positive outcomes, including high levels of patient satisfaction with the care transitions program, significant improvement in participants' confidence with self-care, and actual Medicare cost savings during the 9-month study period of $214,192, excluding the cost to administer the program. The authors concluded that a customized care transitions approach is desirable and often required as the most cost-effective way to manage care transitions and employ evidence-based policy-making. [12]

In a secondary data analysis, Carnahan et al. [13] used the Older Adults Transition Study (OATS) database to identify whether early post-SNF discharge care reduces likelihood of 30-day hospital readmissions in 1,543 patients discharged from a safety-net hospital in Central Indiana to SNF then to home between January 1, 2007 and October 1, 2010. The OATS database combines MDS, Outcome and Assessment Information Set (OASIS), Medicare and Medicaid claims, and electronic medical records. Using a multivariable Cox proportional hazards model adjusting for patient demographic, clinical and utilization variables, the authors found that a home health visit within a week of SNF discharge reduced the hazard of 30-day hospital readmissions [adjusted HR = 0.61, $p < .001$]. Kaplan-Meier survival function estimates found that a home care visit was also significantly associated ($p < 0.05$) with reduced likelihood of readmission at one, two, and three weeks. The authors concluded that early home health services could be a potential intervention to reduce readmissions and improve outcomes for this patient population. [13]

A 2017 systematic review and meta-analysis of 11 randomized controlled trials by Rodakowski et al. [14] examined the effect of integrating informal caregivers (i.e., unpaid individuals who provide support for medical tasks and daily activities once home) into discharge planning from a hospital or SNF on post-discharge readmissions in a combined total sample of 4,361 older adults. The authors reported that integrating caregivers into discharge planning for patients 65 years and older resulted in a 25% lower risk of 90-day hospital readmissions and 24% fewer readmissions at 180 days. These findings provide evidence that community support services can help ensure that patients successfully stay in the community following discharge. [14]

The empirical evidence provided above demonstrates that improvement in successful discharge to community rates among PAC patients is possible through modifying provider-led processes and interventions in the PAC setting and community.

References

2. O'Brien SR, Zhang N. Association between therapy intensity and discharge outcomes in aged Medicare skilled nursing facilities admissions. Archives of Physical Medicine and Rehabilitation. 2018;99(1):107-115.
3. Jung HY, Trivedi AN, Grabowski DC, Mor V. Does more therapy in skilled nursing facilities lead to better outcomes in patients with hip fracture? Physical therapy. 2016;96(1):81-89.
4. Schweickert WD, Pohlman MC, Pohlman AS, et al. Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomised controlled trial. Lancet (London, England). 2009;373(9678):1874-1882.
5. Camicia M, Wang H, DiVita M, Mix J, Niewczyk P. Length of stay at inpatient rehabilitation facility and stroke patient outcomes. Rehabilitation Nursing: The Official Journal of the Association of Rehabilitation Nurses. 2016;41(2):78-90.

6. Thrush A, Rozek M, Dekerlegand JL. The Clinical Utility of the functional status score for the intensive care unit (FSS-ICU) at a long-term acute care hospital: a prospective cohort study. Physical Therapy. 2012;92(12):1536-1545.
7. Kushner DS, Peters KM, Johnson-Greene D. Evaluating Siebens Domain Management Model for inpatient rehabilitation to increase functional independence and discharge rate to home in geriatric patients. Archives of Physical Medicine and Rehabilitation. 2015;96(7):1310-1318.
8. Kushner DS, Peters KM, Johnson-Greene D. Evaluating use of the Siebens Domain Management Model during inpatient rehabilitation to increase functional independence and discharge rate to home in stroke patients. PM&R: The Journal of Injury, Function, and Rehabilitation. 2015;7(4):354-364.
9. Kushner DS, Peters KM, Johnson-Greene D. Evaluating the Siebens Model in geriatric-stroke inpatient rehabilitation to reduce institutionalization and acute-care readmissions. Journal of Stroke and Cerebrovascular Diseases. 2016;25(2):317-326.
10. Berkowitz RE, Jones RN, Rieder R, et al. Improving disposition outcomes for patients in a geriatric skilled nursing facility. Journal of the American Geriatrics Society. 2011;59(6):1130-1136.
11. Buttke D, Cooke V, Abrahamson K, et al. A Statewide Model for assisting nursing home residents to transition successfully to the community. Geriatrics. 2018;3(2):18.
12. Logue MD, Drago J. Evaluation of a modified community based care transitions model to reduce costs and improve outcomes. BMC Geriatrics. 2013;13(1):94.
13. Carnahan JL, Slaven JE, Callahan CM, Tu W, Torke AM. Transitions from skilled nursing facility to home: the relationship of early outpatient care to hospital readmission. Journal of the American Medical Directors Association. 2017;18(10):853-859.
14. Rodakowski J, Rocco PB, Ortiz M, et al. Caregiver integration during discharge planning for older adults to reduce resource use: a metaanalysis. J Am Geriatr Soc. 2017;65(8):1748-1755.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

Successful discharge to community from IRF is widely recognized as an important outcome for patients and their families and an indicator of IRF quality of care.[1-6] The Improving Medicare Post-Acute Care

Transformation Act of 2014 (IMPACT Act of 2014) mandated the specification of a measure to address the domain of discharge to community across post-acute care (PAC) settings.[4] Discharge to community is an important goal for both patients with potential for functional improvement, and patients who may not be expected to make functional gains given their clinical condition or disease. IRFs frequently track their discharge to community outcomes for internal quality improvement purposes; however, this tracking commonly assesses discharge destination alone and not post-discharge readmissions or death. By assessing whether patients successfully remain in the community without acute complications for 31 days following discharge, the DTC-PAC IRF measure is a more meaningful, patient- and family-centered measure of successful community discharge. Discharge to community is an actionable health care outcome as targeted interventions have been shown to successfully increase discharge to community rates in a variety of post-acute settings, including IRFs [7,8] (see attached evidence form). The effectiveness of these interventions suggests that improvement in discharge to community rates among IRF patients is possible through modifying provider-led processes and interventions. In addition to being important from a patient and family perspective, avoiding institutionalization and promoting discharge to community when appropriate are also important from a cost and resource use perspective. Patients discharged to community settings, on average, incur lower costs over the recovery episode, compared with patients discharged to institutional settings.[9-11]

Measuring and publicly reporting the DTC-PAC IRF measure is expected to help differentiate IRFs based on quality of care, and drive improvement in this outcome. The DTC-PAC IRF measure would impact a large number of Medicare beneficiaries and the overall Medicare program. In 2016, Medicare spent $7.7 billion on fee-for service (FFS) IRF care; this care was provided in about 1,200 IRFs nationwide, covering more than 391,000 IRF stays for about 350,000 beneficiaries.[5] There are no NQF-endorsed measures assessing successful discharge to community from IRFs.

[1] Gillsjö C, Schwartz-Barcott D, von Post I. Home: The place the older adult can not imagine living without. BMC Geriatrics. 2011;11(1):10.

[2] Unsworth C. Clients´ perceptions of discharge housing decisions after stroke rehabilitation. American Journal of Occupational Therapy. 1996;50(3):207-216.

[3] Roush CV, Cox JE. The meaning of home: how it shapes the practice of home and hospice care. Home Healthcare Now. 2000;18(6):388-394.

[4] Improving Medicare Post-Acute Care Transformation Act of 2014. Public Law 113–185—October 6, 2014. Available at: https://www.gpo.gov/fdsys/pkg/PLAW-113publ185/pdf/PLAW-113publ185.pdf.

[5] Medicare Payment Advisory Commission. Report to the Congress: Medicare Payment Policy. March 2018. Available at: http://www.medpac.gov/docs/default-source/reports/mar18_medpac_entirereport_sec.pdf.

[6] Hsieh C-H, DeJong G, Groah S, Ballard PH, Horn SD, Tian W. Comparing rehabilitation services and outcomes between older and younger people with spinal cord injury. Archives of Physical Medicine and Rehabilitation. 2013;94(4, Supplement):S175-S186.

[7] Kushner DS, Peters KM, Johnson-Greene D. Evaluating Siebens Domain Management Model for inpatient rehabilitation to increase functional independence and discharge rate to home in geriatric patients. Archives of Physical Medicine and Rehabilitation. 2015;96(7):1310-1318.

[8] Kushner DS, Peters KM, Johnson-Greene D. Evaluating use of the Siebens Domain Management Model during inpatient rehabilitation to increase functional independence and discharge rate to home in stroke patients. PM&R: the Journal of Injury, Function, and Rehabilitation. 2015;7(4):354-364.

[9] Dobrez D, Heinemann AW, Deutsch A, Manheim L, Mallinson T. Impact of Medicare´s prospective payment system for inpatient rehabilitation facilities on stroke patient outcomes. American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists. 2010;89(3):198-204.

[10] Gage B, Morley M, Spain P, Ingber M. Examining Post Acute Care Relationships in an Integrated Hospital System Final Report. RTI International; 2009.

[11] Medicare Payment Advisory Commission. Paying for sequential stays in a unified prospective payment system for post-acute care. In: Report to the Congress: Medicare and the Health Care Delivery System. June 2018. Available at: http://www.medpac.gov/docs/default-source/reports/jun18_ch4_medpacreport_sec.pdf?sfvrsn=0

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

The DTC-PAC IRF measure is based on national data from FY 2016-2017 Medicare FFS inpatient claims. All IRFs paid under Medicare's IRF Prospective Payment System (PPS) and included in the IRF Quality Reporting Program (QRP) are included, provided they had eligible stays. Performance scores were calculated for 1,157 IRFs with eligible stays in FY 2016-2017. Facility-level number of IRF stays ranged from 3 to 6,010 with a mean of 506.0 and standard deviation (SD) of 495.4. Provider characteristics are presented in the testing form (Section 1.5, Table 1). All eligible IRF patient stays discharged between October 1, 2015 and September 30, 2017 were included. For patients with multiple IRF stays in the measurement period, all eligible stays were included. A total of 585,702 patient stays were included in the performance score calculation. Patient demographic and clinical characteristics are presented in the testing form (Section 1.6, Table 2) and attached excel document.

Observed and risk-standardized score distributions for the DTC-PAC IRF measure are presented in the testing form (Table 19 and Figure 4). Observed scores ranged from 29.90% to 100.00%, with a mean of 65.28% and SD of 7.60 percentage points. Risk-standardized performance scores maintained a wide range, from 43.53% to 83.35%, with a mean of 64.74% and SD of 5.36 percentage points.

Observed DTC scores by decile were as follows:

Minimum = 29.90%; 10th percentile (pct) = 55.91%; 20th pct = 59.64%; 30th pct = 61.92%; 40th pct = 63.95%; 50th pct = 65.74%; 60th pct = 67.31%; 70th pct = 69.10%; 80th pct = 71.19%; 90th pct = 74.04%; max = 100.00%.

Risk-standardized DTC scores by decile were as follows:

Minimum = 43.53%; 10th pct = 57.99%; 20th pct = 60.51%; 30th pct = 62.15%; 40th pct = 63.67%; 50th pct = 65.24%; 60th pct = 66.35%; 70th pct = 67.70%; 80th pct = 69.23%; 90th pct = 71.05%; max = 83.35%.

We used bootstrapping to assess the ability of performance measure scores to identify statistically significant differences in provider performance. This analysis was restricted to providers with 25 or more stays during FY 2016-2017 to align with the minimum sample size criterion for public reporting of the measure. Details of the bootstrapping methodology are described in the testing form (Section 2b4.1). We compared the 95% confidence interval of each provider's performance score to the national stay-level observed DTC rate (64.82%) to determine if the provider's performance was significantly different from the national rate. Overall, 95.1% (n = 1,088) of IRFs had performance scores that were significantly different from the national rate, with 44.9% (n = 514) being worse and 50.2% (n = 574) being better than the national rate.

The above variability in performance measure scores, including a range in risk-standardized scores, demonstrates a performance gap and room for improvement in the discharge to community quality domain. Further, the ability of the measure to identify statistically significant differences in scores demonstrates that the measure can discriminate providers based on quality of care.

Given the DTC-PAC IRF measure was recently implemented in October 2016 and has only been publicly reported once, we do not have data to demonstrate improvement in performance over time. In the coming years as more data become available, we will examine score distribution and performance improvement over time.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

The Medicare Payment Advisory Commission (MedPAC) tracks IRFs' risk-adjusted discharge to community and discharge to skilled nursing facility (SNF) rates separately and has reported variation in both these rates across IRFs.[12] A 6 percentage point difference in discharge to community rates and 4.3 percentage point difference in discharge to SNF rates has been reported between the lowest and highest performing quartiles.[12] The MedPAC also reported improvement in discharge to community rates and a slight decrease in discharge to SNF rates between 2011 and 2016. Since the DTC-PAC IRF measure incorporates discharges to SNF, absolute DTC-PAC IRF rates cannot be directly compared to MedPAC rates. Further, some differences in measure exclusion criteria and other specifications make the absolute rates not directly comparable. Nevertheless, the MedPAC data, including presence of a range in risk-adjusted scores, support the existence of a performance gap across IRFs, room for improvement, and improvement over time in the discharge to community domain.

[12] Medicare Payment Advisory Commission. Report to the Congress: Medicare Payment Policy. March 2018. Available at: http://www.medpac.gov/docs/default-source/reports/mar18_medpac_entirereport_sec.pdf

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

The DTC-PAC IRF measure does not adjust for social risk factors to avoid masking potential disparities in quality of care for vulnerable patient groups. It does adjust for age-gender subgroups and original reason for entitlement as these may be associated with the discharge to community outcome for physiological or clinical reasons.

We assessed for disparities in our sample of 1,157 IRFs and 585,702 patient stays using FY 2016-2017 Medicare FFS claims data. In the testing form, we used patient-level dual eligibility for social risk factor testing based on evidence that dual Medicare-Medicaid eligibility is the most important social risk factor predictive of patient outcomes (testing form, Section 2b3.3b). In supplemental analyses (see Appendix B), we also tested race/ethnicity, urbanicity based on beneficiary residence, and socioeconomic status (SES). Socioeconomic status was determined using the Agency of Healthcare Research and Quality's SES Index calculated based on beneficiary residence, with a higher index indicating higher SES. We obtained social risk factor data from the following sources: Integrated Data Repository (IDR) (dual eligibility, race); Inpatient Standard Analytic File (SAF) IRF claim (race when missing in IDR, beneficiary residence ZIP code); Federal Information Processing Standard Publication (FIPS) codes [13] and Rural-Urban Continuum Codes (RUCC_2013) (urbanicity) [14]; and ZIP Code Tabulation Area (ZCTA) and 2016 American Community Survey (5-year file) (AHRQ SES Index). We examined patient-level observed DTC rates across social risk factor subgroups. We also assessed the impact of social risk factors in our logistic regression models.

The distribution of social risk factors across our sample is shown in Appendix Table B-1. A total of 17.6% of our sample was dual eligible, with nearly 12% having full Medicaid benefits. The majority of our sample was white followed by black, and most lived in urban locations.

Some variation in patient-level DTC rates was seen across social risk factor subgroups, with the largest differences based on dual status. The successful DTC rate was 59.5% for duals with full Medicaid, 63.1% for duals without full Medicaid, and 65.7% for non-duals. Black patients had the lowest DTC rates and Asians had the highest rates. There was limited variation in DTC rates based on beneficiary residence location or AHRQ SES Index, with urban location and the first SES index quartile having the lowest DTC rates (Appendix, Table B-2).

In the logistic model testing dual eligibility as the only social risk factor, dual eligibility had a significant negative impact on the DTC outcome, with duals with full Medicaid having a larger negative impact than duals without full Medicaid (testing form, Table 15). Adjusting for dual status increased the model c-statistic by 0.001 only and had a minimal overall impact on facility-level performance scores. The difference between dual-adjusted and non-dual—adjusted scores ranged from -0.87 to 3.57 percentage points, with a mean of 0.00 and standard deviation of 0.36 percentage points (testing form, Table 16).

In the logistic model that included all social risk factors, dual eligibility continued to have the largest negative impact on the DTC outcome, with the estimate for duals with full Medicaid being twice as large as that for duals without full Medicaid (Appendix, Table B-3). Compared with white, all race groups had a protective effective on the DTC outcome, with black having the smallest protective effect and Hispanic having the largest protective effect. Non-urban beneficiary residence had a negative impact on the outcome, with rural residence having a larger negative impact than suburban residence.Unexpectedly, compared with the highest AHRQ SES index quartile, lower quartiles had a protective effect on the outcome with the lowest SES index group (quartile 1) being the most protective. Adjusting for all social risk factors increased the model c-statistic by only 0.001 and had a minimal impact on facility-level performance scores. The difference between social risk factor-adjusted and non-social risk factor—adjusted scores ranged from -1.77 to 2.06 percentage points, with a mean of 0.00 and standard deviation of 0.28 percentage points (Appendix, Table B-4).

[13] https://www.huduser.gov/portal/datasets/usps_crosswalk.html

[14] https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

Not applicable.

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

De.6. Non-Condition Specific*(check all the areas that apply):*

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/Downloads/Measure-Specifications-for-FY17-IRF-QRP-Final-Rule.pdf

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment **Attachment:** NQF_DTC-PAC_IRF_Risk_Adjustment_Logistic_Model_Results_RTI.xlsx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The measure numerator is the risk-adjusted predicted estimate of the number of patients who are discharged to the community, and do not have an unplanned readmission to an acute care hospital or LTCH in the 31-day post-discharge observation window, and who remain alive during the post-discharge observation window.

This estimate starts with the observed number of discharges to community, defined as:

(i) discharges to home or self care with or without home health services, based on Patient Discharge Status Codes 01, 06, 81, or 86 on the Medicare FFS claim [2]; and

(ii) no unplanned acute or LTCH hospitalizations in the 31-day post-discharge window; and

(iii) no death in the 31-day post-discharge window.

The discharge to community outcome is risk-adjusted for patient characteristics and a statistical estimate of the facility effect beyond case-mix (described below).

References

[2] National Uniform Billing Committee Official UB-04 Data Specifications Manual 2018, Version 12, July 2017, Copyright 2017, American Hospital Association.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The numerator uses a model estimated on full national data specific to the IRF setting; it is applied to the IRF's patient stays included in the measure and includes the estimated effect of that IRF. The prediction equation is based on a logistic regression model with a two-level hierarchical structure.

The patient stays in the model have an indicator of the facility they are discharged from; the effect of the facility is measured as a positive or negative shift in the intercept term of the equation. The facility effects are modeled as belonging to a normal (Gaussian) distribution centered at 0 and are estimated along with the effects of patient characteristics in the model.

The risk-adjustment logistic model is re-estimated for every measurement period and model coefficients corresponding to the measurement period are used for measure calculation. Results of the logistic model presented in this submission are based FY 2016-2017 data.

Details about the three components of the measure outcome are described below.

1. DISCHARGE DESTINATION OF COMMUNITY

Discharge to a community destination is determined based on the "Patient Discharge Status Code" from the IRF FFS claim.[3] Discharge to a community destination is defined as discharge to home or self care with or without home health services as described below. While codes 81 and 86 are intended for use on acute care claims only, we observed some instances of these codes on post-acute claims; thus, we include codes 81 and 86 in our community definition.

Discharge Status Codes Indicating Community Discharge:

- 01 = Discharged to home or self care (routine discharge)
- 06 = Discharged/transferred to home under care of organized home health service organization
- 81 = Discharged to home or self care with a planned acute care hospital readmission
- 86 = Discharged/transferred to home under care of organized home health service organization with a planned acute care hospital inpatient readmission

References

[3] National Uniform Billing Committee Official UB-04 Data Specifications Manual 2018, Version 12, July 2017, Copyright 2017, American Hospital Association.

2. UNPLANNED READMISSIONS IN THE 31-DAY POST-DISCHARGE OBSERVATION WINDOW

A patient who is discharged to a community setting is not considered to have a successful discharge to community outcome for this measure if they have a subsequent unplanned readmission to an acute care hospital or LTCH in the post-discharge observation window, which includes the day of discharge and 31 days following day of discharge. We only assess the first readmission encountered in the post-discharge window. Our definition of acute care hospital includes hospitals paid under the Inpatient Prospective Payment System (IPPS), critical access hospitals (CAH), and psychiatric hospitals or units. Using acute care and LTCH claims, we identify unplanned readmissions based on the CMS planned readmissions algorithm[4] used in the following post-acute care (PAC) readmission measures, which have been endorsed by the National Quality Forum (NQF) and used in several CMS programs: (i) Skilled Nursing Facility 30-Day All-Cause Readmission Measure (SNFRM) (NQF #2510); (ii) All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Inpatient Rehabilitation Facilities (IRFs) (NQF #2502); (iii) All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Long-Term Care Hospitals (LTCHs) (NQF #2512); and (iv) Rehospitalization During the First 30 Days of Home Health (NQF #2380).[5][6][7][8] The planned readmission algorithm used in these PAC readmission measures are based on the Hospital-Wide All-Cause Unplanned Readmission Measure (HWR) (NQF #1789)[9], with some additions made for the SNF, IRF, and LTCH setting measures.[10] We used the most current version of the CMS planned readmission algorithm from the 2018 HWR measure specifications for the FY 2016-2017 measure calculation.[4] For future updates, we will use the most current version of the CMS planned readmission algorithm and make necessary updates to the additions made for post-acute care settings to ensure the algorithm corresponds to our measurement period.

The CMS planned readmission definition is based on the claim from the readmission having a code for a diagnosis or procedure that is considered planned; however, if a planned procedure is accompanied by a principal diagnosis in a specified list of acute diagnoses, the readmission is reclassified as unplanned. Readmissions to psychiatric hospitals or units are always classified as planned readmissions.

While the measure was developed with ICD-9-CM procedure and diagnosis codes, it has been transitioned using the ICD-9-CM to ICD-10-CM cross-walk.

References

[4] Appendix E. Planned Readmission Algorithm Version 4.0 2018 (ICD-10). In: 2018 All-Cause Hospital Wide Measure Updates and Specifications Report: Hospital-Level 30-Day Risk-Standardized Readmission Measure – Version 7.0. Available at: https://www.qualitynet.org/dcs/BlobServer?blobkey=id&blobnocache=true&blobwhere=1228890804653&blobheader=multipart%2Foctet-stream&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3Bfilename%3DHospWide_Readmission_AUS_Report_2018_3-28.pdf&blobcol=urldata&blobtable=MungoBlobs or https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1219069855841.

[5] Skilled Nursing Facility 30-Day All-Cause Readmission Measure (SNFRM) (NQF #2510). http://www.qualityforum.org/QPS/2510

[6] All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Inpatient Rehabilitation Facilities (IRFs) (NQF #2502). http://www.qualityforum.org/QPS/2502

[7] All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Long-Term Care Hospitals (LTCHs) (NQF #2512). http://www.qualityforum.org/QPS/2512

[8] Rehospitalization During the First 30 Days of Home Health (NQF #2380). http://www.qualityforum.org/QPS/2380

[9] Hospital-Wide All-Cause Unplanned Readmission Measure (HWR) (NQF #1789). www.qualityforum.org/QPS/1789

[10] Table 2-9. AHRQ CCS Single Level Procedure Codes and ICD-9 Procedure Codes Added to Yale's Planned Readmission Algorithm, for the Post-Acute Care Setting. In: Measure Specifications for Measures Adopted in the FY 2017 IRF QRP Final Rule. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/Downloads/Measure-Specifications-for-FY17-IRF-QRP-Final-Rule.pdf. Note: The ICD-9 codes listed in Table 2-9 were updated with ICD-10-CM codes for data starting October 1, 2015.

3. DEATH IN THE 31-DAY POST-DISCHARGE OBSERVATION WINDOW

A patient who is discharged to a community setting is not considered to have a successful discharge to community outcome for this measure if they die in the post-discharge window, which includes the day of discharge and the 31 days following day of discharge. Death in the post-discharge window is identified based on date of death from Medicare eligibility files.

MEASUREMENT PERIOD

The measure is calculated using two consecutive years of data to ensure adequate number of patient stays for risk-adjustment modeling. All IRF Medicare FFS discharges during the two-year measurement period, except those that meet the exclusion criteria (see S.8. and S.9.), are included in the measure. For patients with multiple stays during the two-year measurement period, each stay is eligible for inclusion in the measure.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

The target population for the measure is the group of Medicare FFS beneficiaries who are discharged from an IRF during the measurement period and are not excluded based on the measure exclusion criteria (see S.8. and S.9.).

The measure denominator is the risk-adjusted expected number of discharges to community. This estimate includes risk-adjustment for patient characteristics with the facility effect removed. The "expected" number of discharges to community is the predicted number of risk-adjusted discharges to community if the same patients were treated at the average facility. The logistic regression model used to calculate the denominator is developed using all non-excluded facility stays in the national data. The denominator is computed in the same way as the numerator, but the facility effect is set at the average.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The measure denominator is the risk-adjusted expected number of discharges to community. See S.8. for details. The target population includes all Medicare FFS beneficiaries who are discharged from an IRF during the measurement period and are not excluded based on the measure exclusion criteria. The target population for the analyses in this submission includes IRF discharges from October 1, 2015 through September 30, 2017 (i.e., FY 2016-2017). Index IRF stays are identified based on discharge date because the Inpatient Standard Analytic File (SAF) from which we extract IRF claims is based on discharge date.

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

Measure exclusion criteria are based on administrative data from Medicare claims and eligibility files. Exclusion criteria were selected to maintain clinical validity of the measure by excluding stays for which discharge to community would not be appropriate, to ensure data availability and completeness, to exclude stays with problematic claims data, and to maintain relevance to the IRF QRP (e.g., excluding IRFs not included in the IRF QRP based on regional location). Only IRF stays that are preceded by a short-term acute care stay in the 30 days prior to the IRF admission date are included in the measure; this is because risk-adjustment variables come from the short-term acute care stay in the 30 days prior to IRF admission. Stays ending in transfers to the same level of care (i.e., IRF-to-IRF discharge) are excluded, because the IRF episode of care had not ended. We also excluded certain discharge status codes on the IRF FFS claim that indicated that the patient was not appropriate for community discharge (e.g., discharges against medical advice). See section S.9 for detailed rationale and data sources for each exclusion.

Measure exclusion criteria are as follows:

- Age under 18 years;
- No short-term acute care hospital discharge within the thirty days preceding an IRF admission;
- Discharges to a psychiatric hospital;
- Discharges against medical advice;
- Discharges to disaster alternative care site or a federal hospital;
- Discharges to court/law enforcement;
- Discharges to hospice or patient stays with a hospice benefit in the 31-day post-discharge window;
- Planned discharges to an acute or LTCH setting;
- Stays for patients without continuous Part A FFS Medicare enrollment during the 12 months prior to the IRF admission date and the 31 days after the IRF discharge;
- IRF stays preceded by a short-term acute care stay for non-surgical treatment of cancer;
- Stays ending in transfer to an IRF;
- Stays with problematic claims data (e.g. anomalous records for stays that overlap wholly or in part or are otherwise erroneous or contradictory; claims not paid);

- Exhaustion of Medicare Part A benefit during the IRF stay; and
- IRF stays in facilities outside of the United States, Puerto Rico, or another U.S. territory.

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

Exclusions for the DTC-PAC IRF measure are listed below, along with the rationale and data source for each exclusion. The measure exclusion criteria are determined by processing Medicare FFS claims and eligibility data to determine whether the individual exclusion criteria are met. All exclusions are based on administrative data.

1. Age under 18 years

Rationale:

a. There is limited literature on discharge destination outcomes in this age group;

b. Patients in this age group represent a different cohort, likely living with their parents, and may be expected to have higher discharge to community rates compared with the rest of the Medicare population; and

c. Patients in this age group represent a small proportion of the IRF Medicare FFS population.

Data source: Birth date and IRF admission date from Inpatient SAF

2. No short-term acute care discharge within the 30 days preceding IRF admission

Rationale: The most recent acute care claim from the 30 days prior to IRF admission provides the principal diagnosis and other important patient data for risk-adjustment. Stays without a short-term acute care discharge within the 30 days prior to PAC admission are excluded because important risk-adjustment data will be missing.

Data source: Hospital discharge date in Inpatient SAF acute care claims in the 30 days before IRF admission

3. Discharges to psychiatric hospital

Rationale: Patients discharged to psychiatric hospital are excluded from the measure because community living at the time of discharge may be potentially inappropriate or unsafe for them due to their mental health or psychiatric condition. This exclusion is also intended to avoid the potential unintended consequence of decreased IRF access for patients discharged from psychiatric hospitals.

Data source: Patient discharge status code from Inpatient SAF IRF claim

4. Discharges against medical advice

Rationale: Stays ending in discharge against medical advice are excluded because the IRF care plan may not have been fully implemented, and the discharge destination may not reflect the facility's discharge recommendation. Additionally, patients discharged against medical advice may potentially be at higher risk of post-discharge readmissions or death, depending on their medical condition, or due to potential non-adherence or non-compliance with care recommendations.

Data source: Patient discharge status code from Inpatient SAF IRF claim

5. Discharges to disaster alternative care sites or federal hospitals

Rationale: Stays ending in discharge to disaster alternative care sites are excluded because these discharges are likely influenced by external emergency conditions and may not represent discretionary discharges by the PAC provider. Discharges to federal hospitals are excluded because we will not have inpatient claims to determine whether the hospitalization was planned or unplanned.

Data source: Patient discharge status code from Inpatient SAF IRF claim

6. Discharges to court/law enforcement

Rationale: Patients who are discharged to court or law enforcement are likely ineligible for discharge to the community due to legal restrictions.

Data source: Patient discharge status code from Inpatient SAF IRF claim

7. Planned discharges to an acute or LTCH setting

Rationale: Planned discharges to an acute care hospital or LTCH are excluded as they indicate that community discharge was not appropriate for the patient. Planned discharges are determined based on the planned readmission algorithm described in section S.5.

Data source: The planned readmission algorithm is applied to diagnosis and procedure codes found on the first Inpatient SAF acute or LTCH claim, if any, on the day of or day after index IRF discharge.

8. Stays ending in discharge to hospice and those with a hospice benefit in the post-discharge observation window

Rationale:

a. Patients discharged to hospice care and those with a hospice benefit in the post-discharge observation window are terminally ill and have very different goals of care compared with non-hospice patients. For non-hospice patients, the primary goal of post-acute care is to return to baseline, independent living in the community; death is an undesirable outcome in the non-hospice population. For hospice patients, the goal is to provide them the opportunity to die comfortably, at home or in a facility.

b. A large proportion of hospice patients die in the 31-day window following discharge from the post-acute setting.

c. The hospice agency, not the post-acute care setting, makes the final decision of discharge to hospice-home or hospice-facility.

Data source: Discharge to hospice is determined based on the Inpatient SAF IRF claim. Post-discharge hospice benefit is determined based on hospice enrollment dates (start and termination dates) in the Enrollment Database (EDB).

9. Patients not continuously enrolled in Part A FFS Medicare for the 12 months prior to the IRF admission date, and at least 31 days after IRF discharge date

Rationale: Patients not continuously enrolled in Medicare Part A FFS for the 12 months prior to IRF admission date are excluded because risk-adjustment for certain comorbidities requires information on acute inpatient bills for one year prior to IRF admission. Patients not continuously enrolled in Medicare Part A FFS for at least 31 days after IRF discharge are excluded because readmissions and death must be observable in the 31-day post-discharge period. Patients without Part A coverage or those who are enrolled in Medicare Advantage plans will not have complete inpatient claims in the system.

Data source: EDB and Denominator Files

10. IRF stays for which the prior short-term acute care stay was for non-surgical treatment of cancer

Rationale: Patient stays for which the prior short-term acute care stay was for non-surgical treatment of cancer are excluded because they have a different trajectory for recovery after discharge, with a high mortality rate. Exclusion of these stays is consistent with the hospital-wide and post-acute readmission measures listed in section S.5.

Data source: Diagnosis codes from the Inpatient SAF prior acute claim

11. IRF stays that end in transfer to the same level of care

Rationale: IRF stays that end in transfer to another IRF are excluded from the measure because the IRF episode has not ended. For an IRF episode that involves transfer to another IRF, only the final IRF provider is included in the measure. (Note that this exclusion does not apply to transitions across different levels of post-acute care (e.g., IRF-to-SNF)).

Data source: Patient discharge status code from Inpatient SAF IRF claim

12. IRF stays with claims data that are problematic (e.g., anomalous records for stays that overlap wholly or in part or are otherwise erroneous or contradictory, stays not matched to the denominator or EDB files, claims not paid)

Rationale: This measure requires accurate information from the post-acute stay and prior short-term acute care stay and from the denominator and EDB files for risk-adjustment.

Data source: Inpatient SAF claims, EDB and denominator files

13. Medicare Part A benefits exhausted

Rationale: Patient stays that have exhausted Medicare Part A coverage during the IRF stay are excluded because the discharge destination decision may be related to exhaustion of benefits.

Data source: Inpatient SAF IRF claim

14. Patient stays from facilities located outside of the United States, Puerto Rico or a U.S. territory

Rationale: Patient stays from foreign facilities may not have complete inpatient claims in the system, and these facilities may not be subject to policy decisions related to this quality measure nor included in the IRF Quality Reporting Program.

Data source: CMS Certification Number from the Inpatient SAF IRF claim

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

Not applicable. Measure is not stratified.

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

Statistical risk model

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

The DTC-PAC IRF measure is risk-adjusted. To develop the risk-adjustment model for this measure, we analyzed Medicare inpatient SAF claims, Denominator, and EDB files, identifying FY 2016-2017 IRF Medicare FFS discharges preceded by an acute care hospitalization (IPPS, CAH, or psychiatric hospital) within 30 days before IRF admission date. We applied the measure exclusion criteria to determine the sample included in the risk-adjustment model. The measure is based on two consecutive fiscal years of data (FY 2016-2017 IRF Medicare FFS discharges).

RISK-ADJUSTMENT VARIABLES

Risk-adjustment variables include demographic and eligibility characteristics; principal diagnoses from the prior short-term acute care stay; IRF case-mix groups (CMG); length of stay, types of surgery or procedures,

and dialysis from the prior short-term acute care stay; comorbidities; and number of prior hospitalizations in the year preceding the IRF admission. See the attached Excel document for the full list of risk-adjusters.

RISK-ADJUSTMENT MODELING AND MEASURE CALCULATION ALGORITHM

We used a hierarchical logistic regression model to predict the probability of discharge to community. Baseline patient characteristics related to the discharge to community outcome and a marker for the specific discharging facility are included in the equation. The equation is hierarchical in that both individual patient characteristics are accounted for, as well as the clustering of patient characteristics by facility. The statistical model estimates both the average predictive effect of the patient characteristics across all facilities, and the degree to which each facility has an effect on discharge to community that differs from that of the average facility. The facility effects are assumed to be randomly distributed around the average (according to a normal distribution).

When computing the facility effect, hierarchical modeling accounts for the known predictors of discharge to community, on average, such as baseline/admission patient characteristics, the observed facility rate, and the number of facility stays eligible for inclusion in the measure. The estimated facility effect is determined mostly by the facility's own data if the number of patient stays is relatively large (as the estimate would be relatively precise) but is adjusted toward the average if the number of stays is small (as that would yield a less precise estimate).

We used the following model:

Let Yij, denote the outcome (equal to 1 if patient i is discharged to community, 0 otherwise) for a patient i at facility j; Zij denotes a set of risk-adjustment variables. We assume the outcome is related to the risk adjusters via a logit function with dispersion:

logit(Prob(Yij =1)) = aj + ß*Zij + eij (1)

aj = µ + ?j ; ?j ~ N(0, t2)

where Zij = (Z1j, Z2j, ... Zkj) is a set of k patient-level risk-adjustment variables; aj represents the facility-specific intercept; µ is the adjusted average outcome across all facilities; t2 is the between-facility variance component; and e ~N(0, s2) is the error term.

The hierarchical logistic regression model is estimated using SAS software (PROC GLIMMIX: SAS/STAT User's Guide, SAS Institute Inc.).

The estimated equation is used twice in the measure. The sum of the probabilities of discharge to community of all patients in the facility measure, including both the effects of patient characteristics and the facility, is the "predicted number" of discharges to community after adjusting for the facility's case mix. The same equation is used without the facility effect to compute the "expected number" of discharges to community for the same patients at the average facility.

The ratio of the predicted-to-expected number of discharges to community (i.e., standardized risk ratio (SRR)) is a measure of the degree to which discharges to community are higher or lower than what would otherwise be expected. The SRR is then multiplied by the national stay-level observed discharge to community rate for all facility stays in the measure, yielding the risk-standardized discharge to community rate for each facility.

The estimation procedure is recalculated for each measurement period. Re-estimating the models for each measurement period allows the estimated effects of the patient characteristics to vary over time as patient case-mix and medical treatment patterns change.

The following steps describe the calculation algorithm/measure logic for the DTC-PAC IRF measure:

Step 1: Identify stays meeting the criteria for the target population, after applying measure exclusions.

Step 2: Identify stays meeting the discharge to community criteria, i.e., discharge to community, no unplanned readmissions on the day of discharge or in the 31 days following discharge, and no death on the day of discharge or in the 31 days following discharge.

Step 3: Identify presence or value of risk-adjustment variables for each patient stay.

Step 4: Calculate the predicted and expected number of discharges to community for each facility using the hierarchical logistic regression model.

The predicted number of discharges to community for each facility (i.e., numerator) is calculated as the sum of the predicted probability of discharge to community for each patient discharged from the facility and included in the measure, including the facility-specific effect.

To calculate the predicted number of discharges to community, predj, for index stays at facilityj, we used the following equation:

predj = Slogit-1(? + ?i + ?*Zij) (2)

where the sum is over all stays in facilityj, and ?i is the facility effect.

The expected number of discharges to community (i.e., denominator) is calculated as the sum of the predicted probability of discharges to community, but without the facility-specific effect included in the predictions. This produces the expected number of discharges at the average facility. To calculate the expected number expj, we used the following equation:

expj = Slogit-1 (? + ?*Zij) (3)

Step 5: Calculate the SRR for each facility, as the ratio of the predicted-to-expected number of discharges to community.

To calculate the facility-level SRR, SRRj, we used the following equation:

SRRj = predj/expj (4)

Step 6: Calculate the risk-standardized discharge to community rate for each facility.

To aid interpretation, the facility-level SRR, SRRj, obtained from equation (4) is then multiplied by the overall national stay-level observed discharge to community rate for all facility stays, ?, to produce the facility-level risk-standardized discharge to community rate (RSRj).

To calculate the risk-standardized discharge to community rate for each facility, we used the following equation:

RSRj = SRRj*? (5)

The DTC-PAC IRF measure is specific to IRF providers only.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

Not applicable. The measure is not based on a sample.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

Not applicable. The measure is not based on a survey or instrument.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Assessment Data, Claims, Management Data

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

This measure is based on IRF Medicare FFS administrative claims and uses data in the Medicare eligibility and inpatient claims files. The eligibility files provide information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A coverage, and periods in the Medicare FFS program. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses, procedures, indicators for use of dialysis services, and indicators of whether the Part A benefit was exhausted. The inpatient claims data files contain patient-level IRF and other hospital records. No data beyond the bills submitted in the normal course of business are required from providers for the calculation of this measure.

The following files are used for specification of the DTC-PAC IRF measure:

MEDICARE INPATIENT SAF for hospital claims, including index IRF claims:

Documentation for the Medicare claims data is provided online by the CMS contractor, Research Data Assistance Center (ResDAC), at the University of Minnesota. The data dictionary for the SAF (also known as Inpatient Research Identifiable File (RIF)) is available at: http://www.resdac.org/cms-data/files/ip-rif/data-documentation.

MEDICARE ENROLLMENT DATABASE:

Information about the EDB may be found at: http://aspe.hhs.gov/datacncl/datadir/cms.htm

MEDICARE DENOMINATOR FILES:

Information and documentation are available at: https://aspe.hhs.gov/report/data-health-and-well-being-american-indians-alaska-natives-and-other-native-americans-data-catalog/medicare-denominator-file and ftp://ftp.cdc.gov/pub/health_statistics/nchs/datalinkage/Denominator%20(edited).pdf

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY (AHRQ) CLINICAL CLASSIFICATIONS SOFTWARE (CCS) GROUPINGS OF ICD-9 AND ICD-10 CODES:

Software for AHRQ CCS groupings is available at https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp (ICD-9) and https://www.hcup-us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp (ICD-10).

AHRQ SURGICAL PROCEDURE CATEGORIES:

These were developed for the HWR measure and are available in SAS programs that are maintained and available upon request.

CMS-HIERARCHICAL CONDITION CATEGORY (HCC) MAPPINGS OF ICD-9 AND ICD-10 CODES:

The full set of CMS-HCC mappings is not currently available publicly. A subset of mappings is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Post-Acute Care

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

Not applicable. This is not a composite performance measure.

2. Validity – See attached Measure Testing Submission Form

NQF_DTC-PAC_IRF_Testing_RTI.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:**
Measure Title: Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRF)
Date of Submission: 8/1/2018

Type of Measure:

☒ **Outcome (*including PRO-PM*)**	☐ **Composite – *STOP – use composite testing form***
☐ Intermediate Clinical Outcome	☐ Cost/resource
☐ Process *(including Appropriate Use)*	☐ Efficiency
☐ 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, 2b1, 2b2, and 2b4 must be completed.**
- **For outcome and resource use measures**, section **2b3** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b5** 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 (2b1-2b6) 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 25 pages (*incuding 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.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Standing 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. For **instrument-based measures** (including PRO-PMs) **and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b1. 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. For **instrument-based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; 12

AND

If patient preference (e.g., informed decision making) 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]

2b3. 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 (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b4. 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.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

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. The degree of consensus and any areas of disagreement must be provided/discussed.

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. 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.17*)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☒ claims	☒ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☒ other: Medicare eligibility data	☒ other: Inpatient Rehabilitation Facility - Patient Assessment Instrument (IRF-PAI) Version 1.3 and Version 1.4; Provider of Services File; Integrated Data Repository

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*).

The Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (DTC-PAC IRF) is based on Medicare fee-for service (FFS) administrative claims and uses data in the Medicare eligibility and inpatient claims files. The eligibility files provide information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A coverage, and periods in the Medicare FFS program. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses, procedures, indicators for use of dialysis services, and indicators of whether the Part A benefit was exhausted. The inpatient claims data files contain stay-level post-acute care (PAC) and other hospital records. No data beyond the bills submitted in the normal course of business are required from providers for the calculation of this measure.

The measure was developed using calendar year 2012-2013 data. This measure submission is based on fiscal year (FY) 2016-2017 data; i.e., IRF discharges from October 1, 2015 through September 30, 2017, which were the most recent data available at the time of our analyses. We used the data sources listed below to develop the claims analytic file for measure specification and testing.

- **Medicare Inpatient Claims Standard Analytical File (SAF) for hospital claims, including index IRF claims:** Documentation for the Inpatient SAF data is provided online by the Centers for Medicare & Medicaid Services (CMS) contractor, Research Data Assistance Center (ResDAC), at the University of Minnesota. The data dictionary for SAF (also known as Inpatient Research Identifiable File (RIF)) is available at: http://www.resdac.org/cms-data/files/ip-rif/data-documentation.
- **Medicare Enrollment Database (EDB):** Information about the EDB is available at: http://aspe.hhs.gov/datacncl/datadir/cms.htm.
- **Medicare Denominator files:** Information and documentation about the Medicare Denominator files are available at:
    - https://aspe.hhs.gov/report/data-health-and-well-being-american-indians-alaska-natives-and-other-native-americans-data-catalog/medicare-denominator-file, and
    - ftp://ftp.cdc.gov/pub/health_statistics/nchs/datalinkage/Denominator%20(edited).pdf
- **Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) groupings of ICD-9 and ICD-10 codes:** Software is available for download at https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp (ICD-9) and https://www.hcup-us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp (ICD-10).
- **CMS-Hierarchical Condition Category (HCC) mappings of ICD-9 and ICD-10 codes:** The full set of CMS-HCC mappings is not currently available publicly. A subset of mappings is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html.

We used three additional data sources for measure testing only, not for specification:

- **Inpatient Rehabilitation Facility-Patient Assessment Instrument (IRF-PAI) Version 1.3 and Version 1.4:**

  The IRF-PAI forms are available at: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/Downloads/IRF-PAI_for_FY2015_final.pdf and https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/Downloads/IRF-PAI-Version-1-4.pdf.
    - We used IRF-PAI items for measure validation, including IRF-PAI discharge destination, falls during the IRF stay, and Section GG discharge functional status items. We used FY 2016-2017 data for IRF-PAI discharge destination. We used FY 2017 data for falls and Section GG discharge functional status items, as these data were only collected beginning October 1, 2016.
- **Beneficiary Fact table (V2_MDCR_BENE_FCT) from the Integrated Data Repository (IDR):**

- We used this data source to extract information on beneficiary dual eligibility status for social risk factor testing. These data are submitted by states to CMS and provide a monthly snapshot representing beneficiary characteristics as of set points in time. We used the BENE_DUAL_STUS_CD (Beneficiary Point Of Sale Dual Status Code) that identifies the entitlement status for the dual eligible beneficiary. General information about the IDR is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/IDR/.

- **Provider of Services Current Files (POS File):**
  - We used this data source to describe the characteristics of IRFs included in specification and testing, such as census region, ownership type, and rurality, reported in **Table 1**. The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the CMS Regional Offices. General information about the POS Files is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html.

1.3. What are the dates of the data used in testing? Fiscal year 2016-2017 (i.e., IRF discharges from October 1, 2015 through September 30, 2017)

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.20*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☐ other:	☐ 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*)

This measure is based on national data. All IRFs paid under Medicare's IRF Prospective Payment System and included in the IRF Quality Reporting Program (QRP) were included, provided they had eligible stays. A total of 1,157 IRFs with eligible stays in FY 2016-2017 were included in measure specification, testing and analysis, with differences noted in section **1.7**. Throughout this form, we use "*k*" to refer to number of providers.

Table 1. Characteristics of IRFs included in Specification and Testing of the FY 2016-2017 DTC-PAC IRF Measure (*k* = 1,157)

Characteristic	%	Characteristic	%
Census Region		**Ownership type**	
New England	3.0	Government	10.6
Mid Atlantic	13.1	For Profit	33.9
East North Central	17.6	Not-For-Profit	38.3
West North Central	8.4	Religious	8.9
South Atlantic	14.1	Other	8.3
East South Atlantic	6.7		
West South Atlantic	20.1	**Rurality**	
Mountain	7.6	Rural	12.9
Pacific	9.0	Urban	87.1
U.S. Territories	0.4		

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: MM201, MM204).

1.6. How many and which <u>patients</u> 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*)

Measure specification and testing was based on national data. All eligible IRF patient stays discharged between October 1, 2015 and September 30, 2017 were included in the measure. For patients with multiple IRF stays in the measurement period, all eligible stays were included. A total of 585,702 patient stays were included in measure specification, testing, and analyses with differences noted in section **1.7**. These eligible stays are referred to as index IRF stays. **Table 2** presents demographic characteristics of the patient stays included in measure specification and testing. Additional characteristics, including but not limited to principal diagnoses, comorbidities, and prior service use are available in the **attached excel document.** Throughout this form, we use "*N*" to refer to number of patient stays.

Table 2. Demographic Characteristics of IRF Patient Stays Included in Specification and Testing of the DTC-PAC IRF Measure (n = 585,702)

Characteristic	%	Characteristic	%
Male	45.5	**Race**	
Female	54.5	White	84.1
Male under 65 years	6.2	Black	10.8
Female under 65 years	5.9	Other	5.2

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: MM204).

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.

The 585,702 patient stays at the 1,157 IRFs described above were included in all aspects of measure testing, with differences described below.

1. **Critical data element validity testing:** Some aspects of data element validity testing (**2b1**) used IRF-PAI discharge assessment data. To do this we matched index IRF claims with IRF-PAI assessments based on patient and facility identification numbers, gender, and admission and discharge dates. A total of 581,811 index IRF stays were successfully matched to IRF-PAI assessments and were included in data element validity testing of the claims discharge status code presented in Section 2b1.3.

Data on falls and functional status were collected on the IRF-PAI discharge assessment beginning October 1, 2016. Thus, data element validity testing using IRF-PAI falls and functional status data could only be conducted on FY 2017 index IRF stays matched to IRF-PAI assessments (n = 290,399). Since the analysis required availability of valid IRF-PAI data, the specific number of index IRF patient stays included in the analyses presented in **Table 5 through Table 8** differs based on IRF-PAI data availability for each item.

2. **Performance measure score testing:**

   Facility-level reliability testing (**2a2**) and examination of statistically significant differences (**2b4**) was restricted to 1,144 IRFs with 25 or more stays in FY 2016-2017; we applied this restriction because this measure will only be publicly reported for providers with a minimum of 25 stays in the two-year measure calculation period.

   Performance measure score empirical validity testing (**2b1**) was conducted on 1,111 IRFs that had FY 2016-2017 data on the three claims-based measures against which DTC scores were correlated.

1.8 What were the social risk factors that were available and analyzed? *For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.*

We used patient-level dual eligibility for social risk factor testing, defining two separate variables: dual eligible with full Medicaid benefits, and dual eligible without full Medicaid benefits. We identified dual eligibility based on the variable BENE_DUAL_STUS_CD extracted from the IDR for the month of IRF admission. The following patient stays were classified as **dual eligible with full Medicaid benefits**: Qualified Medicare beneficiaries (QMB) with full Medicaid, Specified Low-Income Medicare beneficiaries (SLMB) with full Medicaid, and Other dual eligibles with full Medicaid (BENE_DUAL_STUS_CD equal to 02, 04, or 08). We classified the remaining patients with dual status as **dual eligible without full Medicaid benefits** (BENE_DUAL_STUS_CD equals 01, 03, 05, or 06).[1]

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-abstractor 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*)

NQF defines reliability as the repeatability or precision of measurement. It considers reliability of data elements to be the repeatability and reproducibility of the data elements for the same population in the same time period, and reliability of the measure score to be the proportion of variation in the performance scores due to systematic differences across the measured entities (or signal) in relation to random error (or noise).[2]

---

[1] The Centers for Medicare & Medicaid Services. (2017). *Defining Medicare-Medicaid Enrollees in CMS Data Sources (a.k.a. Defining Duals)*. Retrieved from: https://www.cms.gov/Medicare-Medicaid-Coordination/Medicare-and-Medicaid-Coordination/Medicare-Medicaid-Coordination-Office/Downloads/MMCO_DualEligibleDefinition.pdf.

[2] Reliability: NQF's Current Definition of Reliability and Related Concepts – May 16, 2018. Available at: http://www.qualityforum.org/Calendar/2018/05/Scientific_Methods_Panel_In-person_05162018.aspx

1. Critical Data Element Reliability

The DTC-PAC IRF measure is calculated using Medicare FFS claims, which are used for operation of the Medicare program. Since submission of accurate claims is required for provider reimbursement, these data are considered reliable and routinely used in quality measures and for research purposes. Further, CMS uses auditing programs to assess accuracy of claims data. Key data elements used for risk adjustment of this measure including acute care principal diagnosis and procedure codes are used to determine provider payment amount, and thus considered to have high reliability. Since our measure primarily uses risk adjustment variables from the acute hospital claim preceding the IRF stay and hospitalizations in the year preceding the IRF stay, the ability of IRF providers to influence the data to improve their measure performance is minimized. While the outcome is based on the Patient Discharge Status Code on the IRF claim indicating discharge to a community setting, the measure does not rely on this data element alone but confirms discharge to a community destination by looking for acute or long-term care hospital claims on the day of or day after IRF discharge. Presence of such a claim negates the discharge to community identified based on the IRF claim. Section **2b1.1** discusses empirical data element validity testing of the Patient Discharge Status Code on the IRF claim.

The claims data elements used in this measure have been used in several NQF-endorsed measures, further supporting their reliability and appropriateness for use. These measures include the following: (i) Hospital-Wide All-Cause Unplanned Readmission Measure (HWR) (NQF #1789),[3] (ii) Skilled Nursing Facility 30-Day All-Cause Readmission Measure (SNFRM) (NQF #2510); (iii) All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Inpatient Rehabilitation Facilities (IRF) (NQF #2502); (iv) All-Cause Unplanned Readmission Measure for 30 Days post Discharge from Long-Term Care Hospitals (LTCH) (NQF #2512); and (v) Rehospitalization During the First 30 Days of Home Health (NQF #2380).[4, 5, 6, 7]

2. Performance Measure Score Reliability

We conducted two types of performance measure score reliability testing: (i) we examined measure precision by assessing signal-to-noise ratios and the spread of DTC measure scores and confidence intervals; and (ii) we examined a form of measure score repeatability by assessing agreement between a facility's performance measure scores based on randomly-split independent patient stay subsets. Performance-measure score reliability testing was restricted to the 1,144 IRFs that had 25 or more patient stays in the FY 2016-2017 measurement period, as only IRFs with a minimum of 25 patients stays in the two-year measurement period will be publicly reported.

a. Performance Measure Score Precision:

Signal-to-noise ratio is defined as the proportion of variation in performance scores due to systematic differences across the measured entities (i.e., signal) in relation to random error (i.e., noise). It is an indicator of measure precision. We calculated signal-to-noise ratio as the ratio of the facility's performance measure score (signal) to the width of the confidence interval of the performance measure score (noise).

$$Signal\ to\ Noise\ Ratio = \frac{\text{Risk standardized DTC rate}}{\text{Upper 95\% CI of DTC Rate} - \text{Lower 95\% CI of DTC Rate}}$$

[3] Hospital-Wide All-Cause Unplanned Readmission Measure (HWR) (NQF #1789). www.qualityforum.org/QPS/1789

[4] Skilled Nursing Facility 30-Day All-Cause Readmission Measure (SNFRM) (NQF #2510). http://www.qualityforum.org/QPS/2510

[5] All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Inpatient Rehabilitation Facilities (IRF) (NQF #2502). http://www.qualityforum.org/QPS/2502

[6] All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Long-Term Care Hospitals (LTCH) (NQF #2512). http://www.qualityforum.org/QPS/2512

[7] Rehospitalization During the First 30 Days of Home Health (NQF #2380). http://www.qualityforum.org/QPS/2380

Larger signal-to-noise ratios indicate greater precision and reliability; e.g., a ratio of 10 indicates that the signal is 10 times as strong as the noise.

In addition to calculating the signal-to-noise ratio for the overall facility sample of 1,144 IRFs, we stratified the facilities into quartiles based on sample size and calculated signal-to-noise ratios within each quartile. The purpose of this analysis was to assess whether measure precision was acceptable across all providers, irrespective of sample size.

We also examined the spread of the DTC performance measure score and the confidence interval of the measure score.

b. **Split-sample Reliability Testing**: This testing examined agreement between two performance measure scores for a facility based on randomly split, independent subsets of patient stays in the same measurement period. Good agreement between the two performance measure scores calculated in this manner provides evidence that the measure is capturing an attribute of the facility (quality of care) rather than the patient stays (case-mix). We randomly divided each facility's FY 2016-2017 patient stays into halves stratified by fiscal year, thus ensuring that patient stays within each fiscal year were evenly distributed across the split-halves. We calculated performance measure scores for each split-half sample using the same measure specification. We calculated Shrout-Fleiss intraclass correlation coefficients ((ICC (2,1) and ICC (3,1)) between the split-half scores to measure reliability.[8]

   We also calculated ICCs between split-half scores stratified by facility size, to assess whether reliability was acceptable across providers of varying sample size. To do this, we first split our sample of 1,144 IRFs into quartiles based on facility size. We then calculated ICCs within each quartile using the split-half performance measure scores derived above.

2a2.3. For each level of testing 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*)

1. **Critical Data Element Reliability**: Section **2b1.1** discusses critical data element validity testing.
2. **Performance Measure Score Reliability:**
   a. **Measure Precision: Table 3** shows the distribution of the **signal-to-noise ratio** of performance measure scores for the overall sample of 1,144 IRFs included in this testing, and by sample size quartile. The mean signal-to-noise ratio in the overall sample was 77.0 with a range of 23.9 to 92.0. The mean signal-to-noise ratio was lowest in the first sample size quartile and increased progressively with increasing quartile.

      The DTC performance measure score ranged from 43.53 to 83.35 percentage points. The confidence interval width ranged from 0.57 to 1.98 percentage points, with a mean of 0.85 and standard deviation of 0.12 percentage points.

[8] McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, *1*(1), 30.

Table 3. Distribution of provider-level signal-to-noise ratios for the overall IRF sample and by sample size quartile

Facility Sample	*K*	Mean (SD)	Min	25th Pct*	Median	75th Pct	Max
Overall	1,144	77.0 (8.1)	23.9	74.3	78.4	82.0	92.0
25-191 stays (quartile 1)	286	70.9 (10.6)	23.9	67.9	73.6	77.6	86.8
192-337 stays (quartile 2)	285	77.1 (6.4)	33.7	74.7	78.1	80.7	89.4
338-673 stays (quartile 3)	287	79.4 (5.6)	55.2	76.9	79.9	83.0	92.0
674-6,010 stays (quartile 4)	286	80.7 (4.4)	65.3	78.1	81.0	84.0	90.8

* Pct = percentile. RTI Analysis of Medicare Claims File for IRF and LTCH, FY 2016-2017, SNF FY 2017, (reference: BBNQF10).

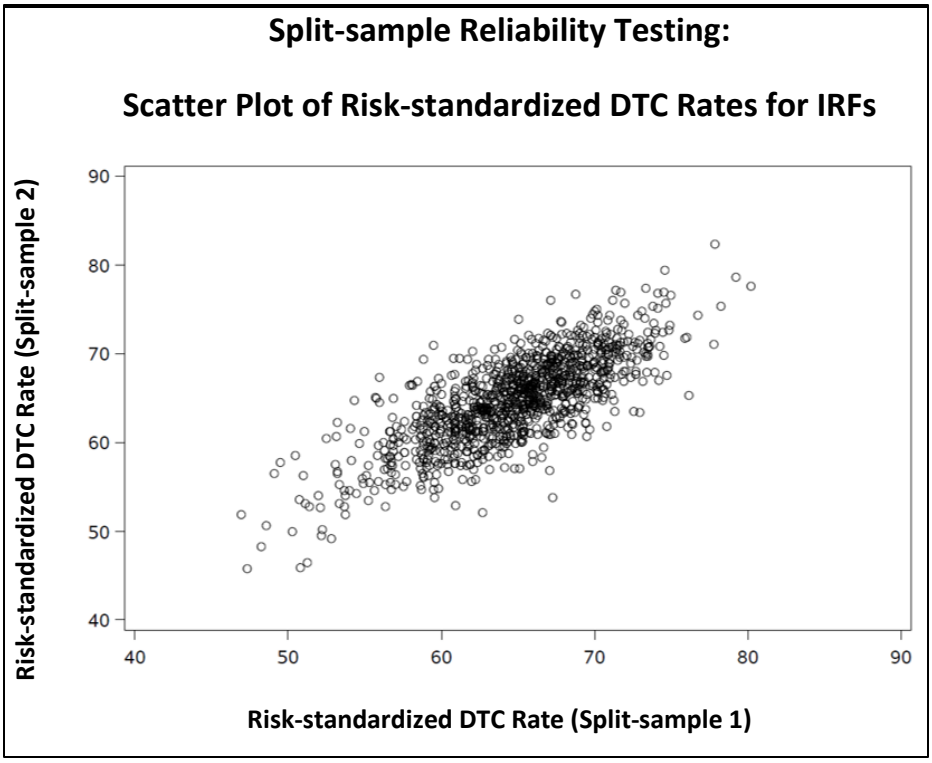
b. **Split-sample Reliability Testing Results: Table 4** presents ICC (2,1) and ICC (3,1) between the split-sample scores for the overall sample of 1,144 IRFs included in this testing, and by sample size quartile. The ICC (2,1) and ICC (3,1) estimates were identical. The ICC in the overall sample was 0.77, with a 95% confidence interval (CI) of 0.74 to 0.79. The ICC was lowest in the first sample size quartile and increased progressively with increasing quartile.

Figures 1a and 1b show scatter plots of the split-half scores overall, and by sample size quartile.

Table 4. Split-sample reliability: Intraclass correlation coefficients between split-sample performance measure scores for the overall IRF sample and by sample size quartile

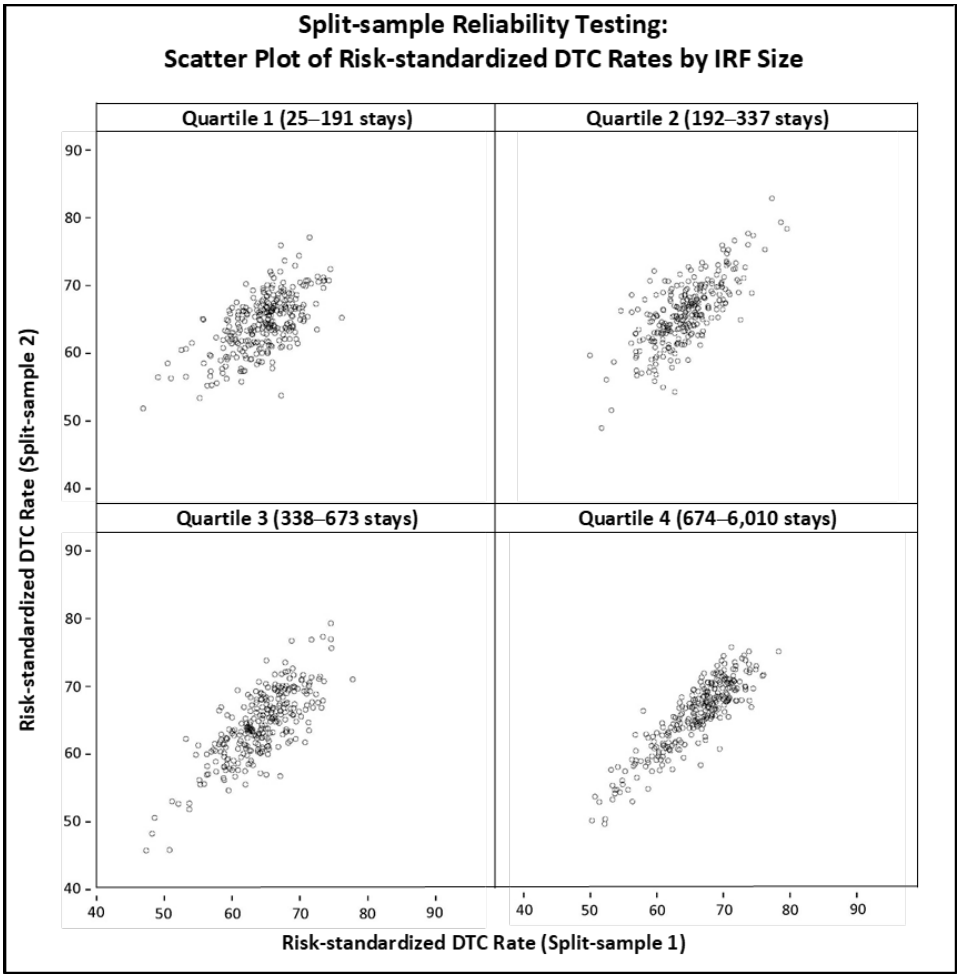
Facility Sample	*K*	ICC (2,1) (95% CI)	ICC (3,1) (95% CI)
Overall	1,144	0.77 (0.74-0.79)	0.77 (0.74-0.79)
25-191 stays (quartile 1)	286	0.62 (0.54-0.69)	0.62 (0.54-0.68)
192-337 stays (quartile 2)	285	0.73 (0.67-0.78)	0.73 (0.67-0.78)
338-673 stays (quartile 3)	287	0.77 (0.62-0.81)	0.77 (0.72-0.81)
674-6,010 stays (quartile 4)	286	0.89 (0.86-0.91)	0.89 (0.86-0.91)

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: MM207a).



RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: MM207a)

Figure 1a. Split-sample Reliability Testing: Scatter Plot of Risk-standardized DTC Rates



RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: MM207a)

Figure 1b. Split-sample Reliability Testing: Scatter Plot of Risk-standardized DTC Rates by IRF Size

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?*)

Overall, reliability testing results indicated good to excellent performance measure score reliability. The mean signal-to-noise ratio was very strong in the overall sample, with the signal being 77 times as strong as the noise, on average. The lowest mean signal-to-noise ratio was 70.9 in the lowest sample size quartile, indicating excellent signal strength even in facilities with the smallest sample sizes. The confidence interval of the performance measure score had a very narrow range, particularly in comparison with the range of the measure scores. These findings indicate that the performance measure score has excellent precision and largely represents systematic differences in quality of care across the measured entities rather than random variation or error.

Intraclass correlation coefficients below 0.5 indicate low reliability; between 0.5 and 0.7, moderate; above 0.7, good; and above 0.9, excellent reliability.[9] The ICC for the overall IRF sample was 0.77, indicating good reliability. When examined by facility sample size, the smallest quartile had moderate reliability, the middle quartiles had good reliability, and the largest quartile had excellent reliability.

2b1. VALIDITY TESTING

---

[9] McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, *1*(1), 30.

2b1.1. 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*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing 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)*

NQF defines validity as the correctness of measurement. It considers validity of data elements to be the correctness of the data elements as compared to an authoritative source, and validity of measure scores to be the correctness of conclusions about quality that can be made based on the measure scores (e.g.., a higher score on a quality measure assessing a positive outcome reflects higher quality).[10]

1. **Critical Data Element Validity Testing:** The discharge to community measure outcome is defined based on three components: (i) discharge to a community setting from the IRF, (ii) no unplanned acute or LTCH readmissions in the 31 days following IRF discharge, and (iii) no death in the 31 days following IRF discharge. The first component, discharge to a community setting, was considered a critical data element for which we conducted validity testing, as it is the first and largest determinant of the discharge to community outcome. The frequency of unplanned readmissions and death in the 31-day post-discharge window is relatively low, making the discharge to community component the key outcome determinant. Discharge to a community setting is defined based on the "Patient Discharge Status Code" reported on the Medicare FFS IRF claim;[11] Discharges to home or self-care with or without home health services, based on Patient Discharge Status Codes 01, 06, 81, or 86 are considered discharge to a community setting.

   We conducted two types of critical data element validity testing: (i) we confirmed correctness of the discharge to community setting component compared to two authoritative sources; and (ii) we examined whether the frequency of discharge to community varies as expected among clinically different patient subgroups expected to have different rates of discharge to a community setting.

   a. **Correctness of discharge to community setting based on IRF claim (critical data element) as compared to discharge to community based on IRF-PAI (authoritative source)**

      This testing included 581,811 index IRF stays that were successfully matched to IRF-PAI assessments. The IRF-PAI discharge to community definition was based on IRF-PAI data element 44D with values of 01-Home (Private home/apt., board/care, assisted living, group home, transitional living) and 06-Home under care of organized home health service organization being classified as discharge to a community setting. Using IRF-PAI as the authoritative source, we examined the percentage of IRF-PAI discharges to community that were also classified as discharged to community based on the IRF claim.

[10] Validity: NQF's Current Definition of Validity and Related Concepts – May 16, 2018. Available at: http://www.qualityforum.org/Calendar/2018/05/Scientific_Methods_Panel_In-person_05162018.aspx

[11] National Uniform Billing Committee Official UB-04 Data Specifications Manual 2018, Version 12, July 2017, Copyright 2017, American Hospital Association.

b. Correctness of discharge to community setting based on IRF claim (critical data element) as confirmed by absence of new inpatient claims on the day of or day after IRF discharge (authoritative source)

This testing was conducted on the total index IRF sample of 585,702 patient stays. For patient stays classified as discharged to a community setting based on the IRF claim, we looked for any acute, LTCH, skilled nursing facility (SNF), or new IRF claims (not duplicative of the index IRF claim) on the day of or day after IRF discharge. Absence of any claims would confirm the discharge to community reported on the index IRF claim. We note that while we also examined for acute, LTCH, SNF or IRF claims in the 2- through 31-day post-discharge window, presence of such claims would not negate the discharge to community setting. The day of and day after post-acute discharge are typically considered the "transfer" period in other claims-based measures and used to determine the patient's discharge destination from the index stay.[12, 13, 14]

c. Known-groups validity testing

Known-groups validity testing evaluates the degree to which a measurement demonstrates different scores for groups known or expected to vary on the construct being measured.[15] We conducted known-groups validity testing by examining whether discharge to community rates based on the IRF claim varied as expected among clinically different patient subgroups expected to have different rates of discharge to a community setting given their clinical status at IRF discharge.

Known-groups validity testing was conducted on the subset of FY 2017 index IRF stays that were successfully matched to IRF-PAI data (n = 290,399). The specific number of stays included in individual analyses presented in **Tables 5-7** differs based on availability of valid IRF-PAI data for each item.

i. Known-groups validity testing for fallers versus non-fallers

We examined discharge to community rates separately for patients who had (i) no fall during the IRF stay; (ii) any fall during the IRF stay; or (iii) a fall with major injury. We used IRF-PAI items J1800 (Any Falls Since Admission) and J1900 (Number of Falls Since Admission) to identify fallers, non-fallers, and fallers with major injury. We hypothesized that patients who had a fall during the IRF stay would have lower discharge to community rates than those who did not have a fall. We further hypothesized that the subgroup of patients who had a fall with major injury would have the lowest discharge to community rates among the three subgroups.

ii. Known-groups validity testing for patients with varying functional abilities with self-care and mobility activities at discharge

We examined the relationship between discharge to community and functional abilities for the seven self-care (GG0130 A-C, GG0130E-H) and the 15 mobility (GG0170 A-G, GG0170I-O) activities at discharge. We used data submitted using the IRF-PAI v1.4 Section GG. We note that for patients who have an incomplete stay or those who refuse assessment, codes indicating "activity not attempted" are entered. We restricted our analysis to patient stays for which functional status

---

[12] Hospital-Wide All-Cause Unplanned Readmission Measure (HWR) (CMS/Yale) (NQF #1789). http://www.qualityforum.org/QPS/1789

[13] All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Inpatient Rehabilitation Facilities (IRF) (NQF #2502). http://www.qualityforum.org/QPS/2502

[14] All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Long-Term Care Hospitals (LTCH) (NQF #2512). http://www.qualityforum.org/QPS/2512

[15] Portney, L. G., & Watkins, M. P. (1993). Foundations of clinical research: application to practice. *Stamford, USA: Appleton & Lange*.

items were assessed at discharge and a functional status rating of 01 (Dependent) through 06 (Independent) was assigned. We hypothesized that patients with greater functional independence levels would be have higher DTC rates.

2. Performance Measure Score Empirical Validity Testing

To assess convergent validity of the performance measure score, we examined whether a facility's DTC score was related as expected to its scores on three other claims-based measures based on the same measurement period: (i) Potentially Preventable 30-Day Post-Discharge Readmission Measure for Inpatient Rehabilitation Facility (IRF) Quality Reporting Program (QRP) (PPR-PD); (ii) Potentially Preventable Within Stay Readmission Measure for Inpatient Rehabilitation Facilities (PPR-WIS); and (iii) Medicare Spending per Beneficiary (MSPB).[16] This testing was conducted on a subset of 1,111 IRFs for which FY 2016-2017 data on these 3 measures were available. We calculated Pearson correlations to assess the association between performance measure scores, and Spearman correlations to assess the association between performance measure score rankings. We hypothesized a negative association between the DTC measure and PPR-PD, PPR-WIS, and MSPB measures, which would indicate that facilities with higher DTC rates have lower potentially preventable readmission rates and lower spending levels. While we considered assessing the relationship of the DTC measure with measures publicly reported on IRF Compare, we found that the measurement periods for the IRF Compare measures were different from that of the DTC measure, thus threatening the validity of the analysis.

3. **Systematic assessment of face validity of the performance measure score as an indicator of quality of care**

As an initial step in measure development, we examined face validity through a systematic and transparent process and by consulting with national experts in the field. We conducted a systematic and comprehensive environmental scan to assess the importance, validity, actionability, feasibility, and burden of the discharge to community measure concept as a quality indicator in addition to IRF-specific literature, we reviewed literature related to SNF and LTCH settings. Further, RTI International and Abt Associates convened a Technical Expert Panel (TEP) comprising experts in quality measurement, clinical care, research, and administration in PAC settings. Three TEP meetings were held during August, September, and October 2015, with one all-day in-person meeting and two webinar meetings. The purpose of these meetings was to get input on the face validity, importance, conceptualization, and draft specifications of the DTC measure. Finally, we solicited public comments on the importance of the DTC-PAC IRF measure via a 30-day public comment period during November-December 2015, and during the FY 2017 IRF PPS rulemaking process.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

1. Critical Data Element Validity

a. Correctness of discharge to community setting based on IRF claim (critical data element) as compared to discharge to community based on IRF-PAI (authoritative source)

Of the 437,031 patient stays classified as discharged to community based on the IRF-PAI, 99.2% were also classified as discharged to the community based on the IRF claim discharge status code.

Table 5. Agreement between Discharge to a Community Setting based on the IRF-PAI and IRF claim

Total N	Discharge to Community Destination		Agreement between IRF-PAI and IRF Claim	
	IRF-PAI	IRF Claim	N	%*
581,811	437,031	434,753	433,532	99.2%

[16] These measures were adopted in the FY 2017 IRF Prospective Payment System (PPS) Final Rule available at: https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf

* The denominator for the percent calculation was the number of IRF-PAI discharges to community. RTI Analysis of Medicare Claims File for IRF and FY 2016-2017 (reference: MM205).

b. Correctness of discharge to community setting based on IRF claim (critical data element) as confirmed by absence of new inpatient claims on the day of or day after IRF discharge (authoritative source)

Only 0.7% of IRF stays that were discharged to a community setting as noted on the IRF claim had an acute, LTCH, SNF, or new IRF claim on the day of or day after index IRF discharge.[17]

c. **Known-groups validity testing for fallers versus non-fallers**

Table 6 presents known-groups testing results for fallers and non-fallers. The successful DTC rate for fallers was 11.4 percentage points lower than that for non-fallers, while the rate for fallers with major injury was 38.7 percentage points lower that for non-fallers.

Table 6. Successful Discharge to Community Rates for Non-Fallers and Non-Fallers

Characteristic	Fallers (n = 18,754)	Fallers with major injury (n = 454)	Non-Fallers (n = 271,645)	Overall Sample (n = 290,399)
Successful DTC	54.4%	27.1%	65.8%	65.1%

RTI Analysis of Medicare Claims File for IRF FY 2017 (reference: MM205).

d. Known-groups validity testing for patients with varying functional independence in self-care and mobility at discharge

Table 7 shows the percentage of patients who were successfully discharged to the community for each functional independence level, ranging from 01 (Dependent) to 06 (Independent), for each self-care activity at discharge. For each self-care activity, a consistent positive monotonic relationship was noted between functional independence level and DTC rates. Discharge to community rates progressively increased with increasing functional independence level. The DTC rate varied greatly by discharge functional status; e.g., only 22.0% of patients who were dependent in eating at discharge were successfully discharged to the community, whereas 75.4% of patients who were independent in eating were successfully discharged to the community.

Tables 8a and 8b show the percentage of patients successfully discharged to the community for each functional independence level for mobility activities at discharge. As noted with the self-care activities, a positive monotonic relationship was noted between functional independence level for mobility activities and successful DTC rates. A large variation in DTC rates by mobility functional independence level was noted. For example, only 13.4% of patients who were dependent in rolling left right at discharge were successfully discharged to the community, whereas 79.3% of patients who were independent were successfully discharged to the community.

[17] RTI Analysis of Medicare Claims File for IRF and LTCH FY 2016-2017 (reference: PP04)

Table 7. Successful Discharge to Community Rates by Self-Care Functional Independence Level at Discharge

Discharge Rating for IRF-PAI Section GG Self-Care Activities	Percent Successfully Discharged to Community						
	Eating	Oral Hygiene	Toilet Hygiene	Shower	Upper Dressing	Lower Dressing	Footwear
01.Dependent	22.0	15.9	20.9	16.9	15.5	20.2	27.8
02. Substantial/maximal assistance	22.6	20.2	33.5	25.5	25.1	32.7	38.0
03. Partial/moderate assistance	26.2	28.5	47.2	45.2	40.9	50.5	55.1
04. Supervision or touching assistance	43.1	47.9	63.1	67.1	55.8	65.4	65.5
05. Setup or clean-up assistance	50.1	50.3	68.1	73.6	59.0	72.1	69.0
06. Independent	75.4	78.7	84.0	85.5	82.5	85.3	84.6

RTI Analysis of Medicare Claims File for IRF FY 2017 (reference: MM205).

Table 8a. Successful Discharge to Community Rates by Mobility Functional Independence Level at Discharge

Discharge Rating for IRF-PAI Section GG Mobility Activities	Percent Successfully Discharged to Community						
	Roll Left Right	Sit to Lying	Lying to Sit	Sit to Stand	Bed to Chair	Toilet Transfer	Car Transfer
01. Dependent	13.4	12.1	12.4	15.4	13.1	15.1	30.4
02. Substantial/maximal assistance	18.4	18.5	18.9	21.3	20.0	22.9	40.7
03. Partial/moderate assistance	33.4	36.6	36.1	35.6	37.6	39.8	56.7
04. Supervision or touching assistance	52.1	54.8	55.1	60.4	61.6	62.7	74.4
05. Setup or clean-up assistance	58.6	60.3	61.0	67.5	68.0	69.6	79.7
06. Independent	79.3	80.9	81.0	84.3	84.9	85.1	86.3

RTI Analysis of Medicare Claims File for IRF FY 2017 (reference: MM205).

Table 8b. Successful Discharge to Community Rates by Mobility Functional Independence Level at Discharge

Discharge Rating for IRF-PAI Section GG Mobility Activities	Percent Successfully Discharged to Community							
	Walk 10 Feet	Walk 50 Feet	Walk 150 Feet	Walk 10 Ft Uneven Surfaces	1 Step	4 Steps	12 Steps	Pick Up Object
01. Dependent	23.1	28.4	40.5	42.0	41.4	44.0	54.5	53.1
02. Substantial/ maximal assistance	28.2	34.6	43.9	44.1	49.6	50.6	60.9	50.6
03. Partial/moderate assistance	40.6	44.2	50.4	57.8	59.9	58.6	65.5	60.7
04. Supervision or touching assistance	63.5	65.6	69.8	74.6	75.2	74.5	78.0	74.2
05. Setup or clean-up assistance	70.7	71.8	74.3	78.3	79.5	79.0	80.3	77.9
06.Independent	85.7	85.9	86.7	86.9	87.2	87.1	87.8	85.1

RTI Analysis of Medicare Claims File for IRF FY 2017 (reference: MM205).

2. Performance Measure Score Empirical Validity Testing

We found a small, significant negative association between DTC measure scores and PPR-PD, PPR-WIS and MSPB measure scores. Both Pearson and Spearman correlations revealed similar relationships (**Table 9**).

Table 9. Relationship between DTC Measure Scores and PPR-PD, PPR-WIS, and MSPB Measure Scores for 1,111 IRFs

Measure	Pearson Correlation		Spearman Correlation	
	Coefficient	p	Coefficient	p
Potentially Preventable 30-Day Post-Discharge Readmissions	-0.11	<0.001	-0.12	<0.001
Potentially Preventable Within Stay Readmissions	-0.22	<0.001	-0.20	<0.001
Medicare Spending per Beneficiary	-0.20	<0.001	-0.16	<0.001

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: MM209).

3. Systematic Assessment of Face Validity:

Note that findings summarized in this section are not statistical results. We found extensive support for the face validity and importance of the discharge to community measure as an indicator of quality of care based on our environmental scan, TEP feedback, and public comment feedback. Discharge to a community setting is an important goal for the majority of IRF patients, making this a high priority measure concept from the patient and family perspective. Home is often considered a symbol of independence, privacy, and competence.[18] Discharge to community is a valuable outcome because it offers a multi-dimensional view of preparation for community life, including cognitive, physical, and psychosocial elements.[19,20,21] In addition to being important from a patient and family perspective, avoiding institutionalization and promoting discharge to community when appropriate, are also important from a cost and resource use perspective. Patients discharged to community settings may incur lower costs over the recovery episode, compared with patients discharged to institutional settings.[22, 23]

[18] Chang PF, Ostir GV, Kuo YF, Granger CV, Ottenbacher KJ. Ethnic differences in discharge destination among older patients with traumatic brain injury. *Archives of physical medicine and rehabilitation.* 2008;89(2):231-236.

[19] El-Solh AA, Saltzman SK, Ramadan FH, Naughton BJ. Validity of an artificial neural network in predicting discharge destination from a postacute geriatric rehabilitation unit. *Archives of physical medicine and rehabilitation.* 2000;81(10):1388-1393.

[20] Tanwir S, Montgomery K, Chari V, Nesathurai S. Stroke rehabilitation: availability of a family member as caregiver and discharge destination. *European journal of physical and rehabilitation medicine.* 2014;50(3):355-362.

[21] Nehra D, Nixon ZA, Lengenfelder C, et al. Acute Rehabilitation after Trauma: Does it Really Matter? *J Am Coll Surg.* 2016;223(6):755-763.

[22] Dobrez D, Heinemann AW, Deutsch A, Manheim L, Mallinson T. Impact of Medicare's prospective payment system for inpatient rehabilitation facilities on stroke patient outcomes. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2010;89(3):198-204.

[23] Gage B, Morley M, Spain P, Ingber M. *Examining Post Acute Care Relationships in an Integrated Hospital System Final Report.* RTI International;2009.

Our TEP summary report is publicly posted.[24] The TEP members unanimously agreed that discharge to community is an important measure concept, noting that return to a community setting is frequently the ultimate goal of rehabilitation. Further, TEP members stated that collecting and reporting DTC rates could be a powerful motivator for providers to work toward facilitating successful return to the community. We also received several public comments during our call for public comment[25] and during the FY 2017 IRF PPS rulemaking process[26] supporting the importance of the DTC-PAC IRF QRP measure as a quality indicator in IRF settings.

2b1.4. 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?*)

1. **Critical Data Element Validity Testing Results**
    a. **Correctness of discharge to community setting based on IRF claim:** Our findings indicated near-perfect accuracy of discharge to community based on the IRF claim as compared with both authoritative sources. Over 99% of patients identified as discharged to community based on the IRF-PAI were also identified as discharged to community based on the IRF claim. Only 0.7% of patients identified as discharged to community based on the IRF claim had an acute, LTCH, SNF or new IRF claim on the day of or day after IRF discharge. The excellent agreement between IRF-PAI and IRF claim discharge to community, along with the absence of institutional transfer claims in 99.3% of patients discharged to community provide strong evidence for critical data element validity of the IRF claim discharge to community component based on the Patient Discharge Status Code.
    b. **Known-groups validity:** Patient-level DTC rates across the different known-groups varied as hypothesized, supporting the ability of this outcome to distinguish between clinically different patient groups. Fallers had lower DTC rates than non-fallers, with fallers with major injury having the lowest DTC rate amongst the three fall subgroups. Patients with progressively higher functional independence in self-care and mobility activities had progressively higher DTC rates. The magnitude of differences in DTC rates across the known-groups for both falls and functional status was large.
2. **Performance Measure Score Validity Testing Results**: As hypothesized, DTC measure scores were negatively associated with the PPR and MSPB measures indicating that facilities with higher DTC rates have lower readmission rates and spending levels. The finding of a small negative correlation coefficient indicates that the DTC measure is associated with the PPR and MSPB measures in the expected direction but is not duplicative and provides unique information about quality of care not captured by these other measures.
3. **Systematic Assessment of Face Validity:** As stated in **2b1.3**, we found extensive support for the face validity and importance of the discharge to community measure as an indicator of quality of care based on our environmental scan, TEP feedback, and public comment feedback.

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section* 2b3

[24] https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Technical-Expert-Panel_-Discharge-to-Community-Quality-Measure.pdf

[25] https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Public-Comment-Summary-Report-for-the-Development-of-a-Discharge-to-Community-Measure.pdf

[26] Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 Federal Register, Vol. 81, No. 151. https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf

2b2.1. 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*)

Measure exclusion criteria and rationale for exclusions are presented in **2b2.3**. We examined the stay-level frequency of each exclusion and the facility-level distribution of exclusions. Most exclusions were required to ensure availability of complete and valid data for measure specification (e.g., exclusion of stays not preceded by a short-term acute care stay in the past 30 days, or exclusion of patients without continuous FFS enrollment), had strong face validity (e.g., discharges against medical advice or to court/law enforcement), or were aligned with exclusion criteria in similar measures (e.g., exclusion of IRF stays preceded by a short-term acute care stay for non-surgical treatment of cancer aligns with the HWR and PAC readmissions measures); thus, we did not examine the impact of these exclusions. We assessed the impact of the hospice exclusion on patient-level outcomes given the possibility that hospice patients may be successfully discharged to the community; our goal was to ensure that the hospice exclusion was supported by the data. We grouped together patients with the two hospice exclusions listed in **Table 10** and examined the frequency of successful DTC and its three components for patients with and without hospice. We hypothesized that hospice patients would have lower successful DTC rates compared with non-hospice patients, warranting their exclusion from the measure.

2b2.2. 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*)

Table 10 displays the overall number and percentage of patient stays excluded based on each criterion. Because the exclusions are not applied sequentially, one patient stay could be excluded for multiple reasons and the sum of individual exclusion frequencies may exceed the total number of patient stays excluded. Overall, 27.1% of stays were excluded because of one or more exclusion criteria. The most frequent reason for exclusion was the absence of a prior acute care stay in the 30 days preceding IRF admission, with 13.8% of being excluded for this reason uniquely. The next most frequent reasons for exclusion were Health Maintenance Organization (HMO) plan enrollment (4.2%), lack of continuous FFS during the required time frame (3.2%), and stays with a hospice benefit in the 31-day post-discharge window (3.1%).

Table 11 shows the distribution of exclusions at the facility-level. On average, 12.8% of a facility's stays were excluded because of absence of a prior acute care stay in the 30 days preceding IRF admission; 4.7% because of HMO enrollment, 3.6% because lack of continuous FFS during the required time frame, and 2.8% because of a hospice benefit in the 31-day post-discharge window.

Table 12 shows a striking difference between hospice and non-hospice patients in the frequency of successful DTC and its three components. Hospice patients had a much lower frequency of discharge to a community setting, and a much higher frequency of unplanned readmissions and death in the 31 days following discharge to community. The overall successful DTC rate was 54.3 percentage points lower in hospice patients compared with non-hospice patients.

Table 10. Stay-Level Frequencies of Exclusion Criteria for the DTC-PAC IRF Measure

Exclusion	N*	%
Age under 18 years	0	0.0%
No short-term acute care discharge within the 30 days preceding an IRF admission as the only reason for exclusion	110,568	13.8%
Discharges to a psychiatric hospital	545	0.1%
Discharges against medical advice	3,442	0.4%
Discharges to federal hospital	157	0.0%
Discharges to disaster alternative care site	3	0.0%
Discharges to court/law enforcement	24	0.0%

Exclusion	N*	%
Discharges to hospice	4,183	0.5%
Patient stays with a hospice benefit in the 31-day post-discharge window;	25,264	3.1%
Stays for patients without continuous Part A FFS Medicare enrollment	25,644	3.2%
Enrolled in an HMO	33,500	4.2%
Managed Care Paid for Claim	744	0.1%
IRF stays preceded by a short-term acute care stay for non-surgical treatment of cancer	4,907	0.6%
Stays ending in transfer to an IRF	1,682	0.2%
Planned discharges to an acute or LTCH setting	3,603	0.4%
Stays with Problematic Claim Data	14055	1.8%
Exhaustion of Medicare Part A benefit during the IRF stay	106	0.0%
IRF stays in facilities outside of the United States, Puerto Rico, or another U.S. territory	0	0.0%
Total Number of Patient-Stays Excluded	**217,203**	**27.1%**

*Exclusions are not mutually exclusive; one patient-stay could be excluded for multiple reasons. The sum of individual exclusion frequencies may exceed the total number of patients excluded.

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: BBNQF12)

Table 11. Facility-Level Distribution of Exclusion Criteria for the DTC-PAC IRF Measure

Exclusion	Mean	Min	25th	Median	75th	Max
Age under 18 years	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
No short-term acute care discharge within the 30 days preceding an IRF admission as the only reason for exclusion	12.8%	0.0%	6.7%	10.0%	16.1%	100.0%
Discharges to a psychiatric hospital	0.1%	0.0%	0.0%	0.0%	0.0%	11.1%
Discharges against medical advice	0.4%	0.0%	0.0%	0.3%	0.6%	7.7%
Discharges to federal hospital	0.0%	0.0%	0.0%	0.0%	0.0%	2.0%
Discharges to disaster alternative care site	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%
Discharges to court/law enforcement	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%
Discharges to hospice	0.5%	0.0%	0.0%	0.4%	0.7%	4.1%
Patient stays with a hospice benefit in the 31-day post-discharge window	2.8%	0.0%	1.7%	2.6%	3.7%	9.9%
Stays for patients without continuous Part A FFS Medicare enrollment	3.6%	0.0%	2.3%	3.1%	4.4%	100.0%
Enrolled in an HMO	4.7%	0.0%	2.1%	3.3%	5.0%	68.4%
Managed Care Paid for Claim	0.1%	0.0%	0.0%	0.0%	0.1%	4.4%
IRF stays preceded by a short-term acute care stay for non-surgical treatment of cancer	0.6%	0.0%	0.2%	0.5%	0.9%	6.5%
Stays ending in transfer to an IRF	0.4%	0.0%	0.0%	0.1%	0.3%	100.0%
Planned discharges to an acute or LTCH setting	0.5%	0.0%	0.1%	0.4%	0.7%	9.1%
Stays with Problematic Claim Data	1.9%	0.0%	0.4%	1.1%	2.4%	100.0%
Exhaustion of Medicare Part A benefit during the IRF stay	0.0%	0.0%	0.0%	0.0%	0.0%	1.3%
IRF stays in facilities outside of the United States, Puerto Rico, or another U.S. territory	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: BBNQF12)

Table 12. Impact of the Hospice Exclusion on Patient-Level DTC Outcomes

Variable	Non-hospice (N = 585,702)	Hospice (N = 18,942)
Discharge to community setting (codes 1, 6, 81, 86)	74.7%	34.5%
Unplanned readmissions following return to community	12.8%	49.2%
Death following return to community	1.0%	47.9%
Successful DTC	**64.8%**	**10.5%**

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: BBNQF12)

2b2.3. 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*)

Exclusions for the discharge to community measure are listed in **Table 13**, along with the rationale for each exclusion. While the no prior acute stay exclusion results in exclusion of 13.8% of patient stays, this exclusion is essential because key risk adjustment data are obtained from the prior acute care claim. The other most frequent exclusions including HMO enrollment and lack of continuous FFS enrollment are required to ensure availability of complete data during the index IRF stay, the year preceding IRF admission and 31 days following discharge from IRF, as the measure is specified using data from this entire time frame. Analysis of the impact of the hospice exclusion on patient-level DTC outcomes provides strong empirical support for this exclusion as hospice patients have strikingly lower DTC rates compared with non-hospice patients.

Table 13. Exclusion Criteria and Rationale for the DTC-PAC IRF Measure

Exclusion	Rationale
Age under 18 years	There is limited literature on discharge destination outcomes in this age group. Patients in this age group represent a different cohort, likely living with their parents, and may be expected to have higher discharge to community rates compared with the rest of the Medicare population. Patients in this age group represent a small proportion of the post-acute Medicare FFS population.
No short-term acute care discharge within the 30 days preceding IRF admission	Acute care claims from the 30 days prior to IRF admission provide the principal diagnosis and other important patient data for risk adjustment. Stays without a short-term acute care discharge within the 30 days prior to IRF admission are excluded because important risk adjustment data will be missing.
Discharges to psychiatric hospital	Patients discharged to psychiatric hospital are excluded from the measure because community living at the time of discharge may be potentially inappropriate or unsafe for them due to their mental health or psychiatric condition.
Discharges against medical advice	Stays ending in discharge against medical advice are excluded because the IRF care plan may not have been fully implemented, and the discharge destination may not reflect the facility's discharge recommendation. Additionally, patients discharged against medical advice may potentially be at higher risk of post-discharge readmissions or death, depending on their medical condition, or due to potential non-adherence or non-compliance with care recommendations.

Exclusion	Rationale
Discharges to disaster alternative care sites or federal hospitals	Stays ending in discharge to disaster alternative care sites are excluded because these discharges are likely influenced by external emergency conditions and may not represent discretionary discharges by the IRF provider. Discharges to federal hospitals are excluded because we will not have inpatient claims to determine whether the hospitalization was planned or unplanned.
Discharges to court/law enforcement	Patients who are discharged to court or law enforcement are likely ineligible for discharge to the community due to legal restrictions.
Planned discharges to an acute or LTCH setting	Planned discharges to an acute care hospital or LTCH are excluded as they indicate that community discharge was not appropriate for the patient. Planned discharges are determined based on the planned readmission algorithm.[27]
Stays ending in discharge to hospice and those with a hospice benefit in the post-discharge observation window	Patients discharged to hospice care and those with a hospice benefit in the post-discharge observation window are terminally ill and have very different goals of care compared with non-hospice patients. For non-hospice patients, the primary goal of post-acute care is to return to baseline, independent living in the community; death is an undesirable outcome in the non-hospice population. For hospice patients, the goal is to provide them the opportunity to die comfortably, at home or in a facility. A large proportion of hospice patients die in the 31-day window following discharge from the post-acute setting. The hospice agency, not the post-acute care setting, makes the final decision of discharge to hospice-home or hospice-facility.
Patients not continuously enrolled in Part A FFS Medicare for the 12 months prior to the IRF admission date, and at least 31 days after IRF discharge date	Patients not continuously enrolled in Medicare Part A FFS for the 12 months prior to IRF admission date are excluded because risk adjustment for certain comorbidities requires information on acute inpatient bills for one year prior to IRF admission. Patients not continuously enrolled in Medicare Part A FFS for at least 31 days after IRF discharge are excluded because readmissions and death must be observable in the 31-day post-discharge period. Patients without Part A coverage or those who are enrolled in Medicare Advantage plans will not have complete inpatient claims in the system.
IRF stays for which the prior short-term acute care stay was for non-surgical treatment of cancer	Patient stays for which the prior short-term acute care stay was for non-surgical treatment of cancer are excluded because they have a different trajectory for recovery after discharge, with a high mortality rate. Exclusion of these stays is consistent with the hospital-wide and post-acute readmission measures.

[27] Appendix E. Planned Readmission Algorithm Version 4.0 2018 (ICD-10). In: 2018 All-Cause Hospital Wide Measure Updates and Specifications Report: Hospital-Level 30-Day Risk-Standardized Readmission Measure – Version 7.0. Available at: https://www.qualitynet.org/dcs/BlobServer?blobkey=id&blobnocache=true&blobwhere=1228890804653&blobheader=multipart%2Foctet-stream&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3Bfilename%3DHospWide_Readmission_AUS_Report_2018_3-28.pdf&blobcol=urldata&blobtable=MungoBlobs or https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1219069855841.

Exclusion	Rationale
IRF stays that end in transfer to the same level of care	IRF stays that end in transfer to another IRF are excluded from the measure because the IRF episode has not ended. For an IRF episode that involves transfer to another IRF, only the final IRF provider is included in the measure. (Note that this exclusion does not apply to transitions across different levels of post-acute care (e.g., IRF to SNF)).
IRF stays with claims data that are problematic (e.g., anomalous records for stays that overlap wholly or in part or are otherwise erroneous or contradictory, stays not matched to the denominator or EDB files, claims not paid)	This measure requires accurate information from the post-acute stay and prior short-term acute care stay and from the denominator and EDB files for risk adjustment.
Medicare Part A benefits exhausted	Patient stays that have exhausted Medicare Part A coverage during the IRF stay are excluded because the discharge destination decision may be related to exhaustion of benefits.
Patient stays from facilities located outside of the United States, Puerto Rico or a U.S. territory	Patient stays from foreign facilities may not have complete inpatient claims in the system, and these facilities may not be subject to policy decisions related to this quality measure nor included in the IRF Quality Reporting Program.

2b3. 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 2b4**.**

2b3.1. What method of controlling for differences in case mix is used?

☐ **No risk adjustment or stratification**

☒ **Statistical risk model with** 228 **risk factors**

☐ **Stratification by risk categories**

☐ **Other,**

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

Risk adjustment methodology

We used a hierarchical logistic regression model to predict the probability of discharge to community. Baseline patient characteristics related to the discharge to community outcome and a marker for the specific discharging facility are included in the equation. The equation is hierarchical in that both individual patient characteristics are accounted for, as well as the clustering of patient characteristics by facility. The statistical model estimates both the average predictive effect of the patient characteristics across all facilities, and the degree to which each facility has an effect on discharge to community that differs from that of the average facility. The facility effects are assumed to be randomly distributed around the average (according to a normal distribution).

When computing the facility effect, hierarchical modeling accounts for the known predictors of discharge to community, on average, such as baseline/admission patient characteristics, the observed facility rate, and the number of facility stays eligible for inclusion in the measure. The estimated facility effect is determined mostly by the facility's own data if the number of patient stays is relatively large (as the estimate would be relatively precise) but is adjusted toward the average if the number of stays is small (as that would yield a less precise estimate).

We used the following model:

Let Y_{ij}, denote the outcome (equal to 1 if patient i is discharged to community, 0 otherwise) for a patient i at facility j; Z_{ij} denotes a set of risk adjustment variables. We assume the outcome is related to the risk adjusters via a logit function with dispersion:

$$\text{logit}(\text{Prob}(Y_{ij}=1)) = \alpha_j + \beta * Z_{ij} + \varepsilon_{ij} \quad (1)$$

$$\alpha_j = \mu + \omega_j ; \ \omega_j \sim N(0, \tau^2)$$

where $Z_{ij} = (Z_{1j}, Z_{2j}, \ldots Z_{kj})$ is a set of k patient-level risk adjustment variables; α_j represents the facility-specific intercept; μ is the adjusted average outcome across all facilities; τ^2 is the between-facility variance component; and $\varepsilon \sim N(0, \sigma^2)$ is the error term.

The hierarchical logistic regression model is estimated using SAS software (PROC GLIMMIX: SAS/STAT User's Guide, SAS Institute Inc.).

The estimated equation is used twice in the measure. The sum of the probabilities of discharge to community of all patients in the facility measure, including both the effects of patient characteristics and the facility, is the "predicted number" of discharges to community after adjusting for the facility's case mix. The same equation is used without the facility effect to compute the "expected number" of discharges to community for the same patients at the average facility.

The ratio of the predicted-to-expected number of discharges to community (i.e., standardized risk ratio (SRR)) is a measure of the degree to which discharges to community are higher or lower than what would otherwise be expected. The SRR is then multiplied by the national stay-level observed discharge to community rate for all facility stays in the measure, yielding the risk-standardized discharge to community rate for each facility.

The estimation procedure is recalculated for each measurement period. Re-estimating the models for each measurement period allows the estimated effects of the patient characteristics to vary over time as patient case-mix and medical treatment patterns change.

Risk Adjustment Variables

Risk adjustment variables include demographic and eligibility characteristics; principal diagnoses from the prior short-term acute care stay; IRF case-mix groups (CMG); length of stay, types of surgery or procedures, and dialysis from the prior short-term acute care stay; comorbidities; and number of prior hospitalizations in the year preceding the IRF admission. See the **attached Excel document** for the full list of risk-adjusters. Principal diagnosis codes from the prior acute claim are grouped clinically using AHRQ CCS groupings for ICD codes referenced in section **1.2**. Surgical procedure categories, if present, are based on procedure codes from the prior acute claim in the past 30 days, as defined in the HWR measure. The procedures are grouped using the CCS classification developed by AHRQ.[28] Comorbidities are based on secondary diagnoses from the prior acute claim in the past 30 days alone, or along with all diagnosis codes from additional prior acute claims in the year preceding IRF admission. These diagnosis codes are clustered using the CMS-HCC groupings; we used HCC version 22 for the FY 2016-2017 measure specification. The decision to use the comorbidities based on the prior 30-day acute care hospitalization versus a one-year lookback was based on statistical coefficient results and prior clinical input we received during development of the IRF all-cause readmission measure in terms of which types of comorbidities are acute exacerbations and more temporary, versus conditions that are more chronic and long-term.

While the measure was developed with ICD-9-CM procedure and diagnosis codes, it has been transitioned using the ICD-9-CM to ICD-10-CM cross-walk.

Measure Calculation Algorithm

The following steps describe the calculation algorithm for the DTC-PAC IRF measure:

[28] These were developed for the Hospital-Wide All-Cause Unplanned Readmission measure and are available in SAS programs that are maintained and available upon request.

Step 1: Identify stays meeting the criteria for the target population, after applying measure exclusions.

Step 2: Identify stays meeting the discharge to community criteria, i.e., discharge to community, no unplanned readmissions on the day of discharge or in the 31 days following discharge, and no death on the day of discharge or in the 31 days following discharge.

Step 3: Identify presence or value of risk adjustment variables for each patient stay.

Step 4: Calculate the predicted and expected number of discharges to community for each facility using the hierarchical logistic regression model.

The predicted number of discharges to community for each facility (i.e., numerator) is calculated as the sum of the predicted probability of discharge to community for each patient discharged from the facility and included in the measure, including the facility-specific effect.

To calculate the predicted number of discharges to community, $pred_j$, for index facility stays at $facility_j$, we used the following equation:

$$pred_j = \Sigma logit^{-1}(\mu + \omega_j + \beta * Z_{ij}) \quad (1)$$

where the sum is over all stays in $facility_j$, and ω_j is the facility effect.

The expected number of discharges to community (i.e., denominator) is calculated as the sum of the predicted probability of discharges to community, but without the facility-specific effect, ω_j, included in the predictions. This produces the expected number of discharges at the average facility. To calculate the expected number exp_j, we used the following equation:

$$exp_j = \Sigma logit^{-1}(\mu + \beta * Z_{ij}) \quad (2)$$

Step 5: Calculate the SRR for each facility, as the ratio of the predicted to expected number of discharges to community.

To calculate the facility-level SRR, SRR_j, we used the following equation:

$$SRR_j = pred_j / exp_j \quad (3)$$

Step 6: Calculate the risk-standardized discharge to community rate for each facility.

To aid interpretation, the facility-level SRR, SRR_j, obtained from equation (4) is then multiplied by the overall national observed discharge to community rate for all facility stays, $\bar{Y}$, to produce the facility-level risk-standardized discharge to community rate (RSR_j).

To calculate the risk-standardized discharge to community rate for each facility, we used the following equation:

$$RSR_j = SRR_j * \bar{Y} \quad (4)$$

The DTC-PAC IRF measure is specific to IRF providers only.

2b3.2. If an outcome or resource use component measure is <u>not risk adjusted or stratified</u>, provide <u>rationale and analyses</u> to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable.

2b3.3a. Describe the conceptual/clinical <u>and</u> statistical methods and criteria used to select patient factors (clinical factors or social risk 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) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

Conceptual rationale for risk factor selection

The goal of risk adjustment is to account for differences across facilities in patient demographic and clinical characteristics that might be related to the outcome but exist at or prior to the IRF admission, such that differences in performance measure scores can be attributed to differences in quality of care rather than case-mix. Our model includes clinical characteristics present at IRF admission and during the prior acute hospitalization in the past 30 days, and certain comorbidities from acute care hospitalizations in the year preceding IRF admission. To develop our risk adjustment model, we built on extensive work conducted during risk adjustment model development for the IRF all-cause readmission measure, since both measures use the same data sources and examine post-discharge outcomes over similar periods. We tested the list of risk adjusters from the IRF all-cause readmission measure and made modifications or additions as needed based on statistical findings.[29] In addition to principal diagnoses from the prior acute stay, we adjust for IRF Case-Mix Groups (CMGs). CMGs are mutually exclusive groupings built on the patient's underlying medical problem related to IRF care, functional motor score, cognitive score or age. This variable adds valuable predictive information beyond the prior acute care diagnosis.

Variables such as race or ethnicity are often not included as adjustment variables in models because the standards of care should not vary across demographic markers for vulnerability to disparities in health outcomes and receipt of quality care. However, for some outcomes, an argument can be made that some of these markers (sex and age) are also associated with demonstrated clinical/physiologic differences that can determine risk at the time the patient enters the IRF; thus, we adjust for age-sex groups in our model.

Several risk adjusters used in our model were supported by our environmental scan findings which spanned all inpatient PAC settings. Age has frequently been found to be associated with discharge destination, with younger patients being more likely to be discharged to the community.[30,31,32] An association between gender and discharge to community outcomes has also been reported, although there is lack of consensus about whether men or women are more likely to be discharged home.[33,34,35] Certain diagnoses including diabetes, stroke, Alzheimer's disease or dementia, psychiatric disorders, cancer, and end-stage renal disease, are

---

[29] All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Inpatient Rehabilitation Facilities (IRFs) (CMS/RTI) (NQF# 2502), Available at: http://www.qualityforum.org/QPS/2502

[30] Reistetter TA, Graham JE, Deutsch A, Granger CV, Markello S, Ottenbacher KJ. Utility of functional status for classifying community versus institutional discharges after inpatient rehabilitation for stroke. *Archives of physical medicine and rehabilitation.* 2010;91(3):345-350.

[31] Corrigan JD, Horn SD, Barrett RS, et al. Effects of Patient Preinjury and Injury Characteristics on Acute Rehabilitation Outcomes for Traumatic Brain Injury. *Archives of Physical Medicine and Rehabilitation.* 2015;96(8 Suppl):S209-221.e206.

[32] Arling G, Kane RL, Cooke V, Lewis T. Targeting residents for transitions from nursing home to community. *Health services research.* 2010;45(3):691-711.

[33] Chang PF, Ostir GV, Kuo YF, Granger CV, Ottenbacher KJ. Ethnic differences in discharge destination among older patients with traumatic brain injury. *Archives of physical medicine and rehabilitation.* 2008;89(2):231-236.

[34] Leland NE, Gozalo P, Christian TJ, et al. An Examination of the First 30 Days After Patients are Discharged to the Community From Hip Fracture Postacute Care. *Medical Care.* 2015;53(10):879-887

[35] Berges IM, Kuo YF, Ostir GV, Granger CV, Graham JE, Ottenbacher KJ. Gender and ethnic differences in rehabilitation outcomes after hip-replacement surgery. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2008;87(7):567-572.

associated with a decreased likelihood of discharge to community.[36,37,38,39] Presence of comorbidities and severity of comorbidities has also been correlated with a decreased likelihood of discharge to community.[40,41,42]

Statistical approach for risk factor selection

Our initial risk adjustment models were developed using logistic regression, reducing the need for greater computational intensity of hierarchical modeling while model building. We developed the initial model using CY 2012-2013 data and re-estimated and refined the model using FY 2016-2017 data. While there are no conceptual differences between the models for the two measurement periods, the specific list of risk adjusters varies slightly based on statistical findings in the current data. Since this is a predictive within sample rather than hypothesis-testing model and the goal is to account for the effect of case-mix to allow fair provider comparisons, we did not apply a strict alpha criterion of 0.05 for statistical significance; we wanted our model to account for all differences in case-mix to allow a fair comparison of providers' quality of care. In general, we used a *p*-value of 0.4 as a criterion for keeping variables in the model; however, we retained variables with higher *p*-values if they were thought to be clinically important or statistically important based on estimate sizes. We excluded risk factors, such as comorbidities, that had a clinically unexpected protective effect and may have reflected coding practices rather than a patient's clinical status. For example, we hypothesized that certain comorbidities that were protective may have been coded more often for healthier patients with fewer severe comorbidities than for sicker patients who had more competing comorbidities to include in the billing form. While we did not have a strict minimum sample size criterion for each covariate, we examined sample sizes and combined clinically similar covariates with small samples to improve statistical power and stability of estimates.

Our final model included 228 risk adjusters; among these were 36 aggregate CCS groupings and 28 individual CCS groupings, 28 CMG groupings and 3 individual CMGs, 82 individual HCCs, and 12 HCC clusters. 50 HCCs were based on secondary diagnoses from the prior acute stay and the remaining 44 HCCs were based on secondary diagnoses from the prior acute stay as well as diagnoses obtained from additional acute care claims in the 365 days preceding IRF admission.

---

[36] Arling G, Kane RL, Cooke V, Lewis T. Targeting residents for transitions from nursing home to community. *Health services research*. 2010;45(3):691-711.

[37] Arling G, Williams AR, Kopp D. Therapy Use and Discharge Outcomes for Elderly Nursing Home Residents. *Gerontologist*. 2000;40(5):587-595.

[38] Thakar CV, Quate-Operacz M, Leonard AC, Eckman MH. Outcomes of hemodialysis patients in a long-term care hospital setting: a single-center study. *American journal of kidney diseases* : the official journal of the National Kidney Foundation. 2010;55(2):300-306

[39] Jackson JP, Whisner S, Wang EW. A predictor model for discharge destination in inpatient rehabilitation patients. *American journal of physical medicine & rehabilitation* / Association of Academic Physiatrists. 2013;92(4):343-350.

[40] Wang H, Niewczyk P, Divita M, et al. Impact of pressure ulcers on outcomes in inpatient rehabilitation facilities. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists.* 2014;93(3):207-216.

[41] Roberts PS, DiVita MA, Riggs RV, Niewczyk P, Bergquist B, Granger CV. Risk Factors for Discharge to an Acute Care Hospital From Inpatient Rehabilitation Among Stroke Patients. *PM&R.* 2014;6(1):50-55

[42] Spruit-Van Eijk M, Zuidema SU, Buijck BI, Koopmans RTcm, Geurts ACH. Determinants of rehabilitation outcome in geriatric patients admitted to skilled nursing facilities after stroke: a Dutch multi-centre cohort study. *Age & Ageing.* 2012;41(6):746-752

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☒ **Internal data analysis**

☐ **Other (please describe)**

Given there is limited literature on the relationship between social risk factors and discharge to community outcomes, we examined the literature on social risk factors in similar claims-based readmission measures. Since all-cause unplanned readmissions are a component of the DTC-PAC IRF measure outcome, literature related to readmissions measures is relevant to this measure. We selected patient-level dual eligibility status for social risk factor testing based on extensive evidence from several readmissions measures that dual eligibility is the most important social risk factor predictive of patient outcomes. Dual eligibility data also have the advantage of being easily available and more patient specific than area-level variables. The NQF sociodemographic trial period for the IRF all-cause readmission measures (NQF#2502), conducted by our team, is the most directly related evidence regarding the importance of dual eligibility compared to other social risk factors.[43,44] In the analysis we conducted for NQF's sociodemographic trial period, we examined both patient-level social risk factors (race/ethnicity and Medicare/Medicaid dual status) and 21 county-level social risk factors such as median household income, median home value, percent of people under certain poverty thresholds, and number of primary care physicians. We found that patient-level information on dual eligibility was significantly associated with higher odds of readmission but found inconsistent results for the county-level risk factors. For example, contrary to our hypothesis, we found that higher percent of residents in a county with less than a high school diploma and higher percent of residents not speaking English were associated with lower odds of readmission, while higher county-level median household income was associated with higher odds of readmission. Further, starting in fiscal year 2019, CMS's Hospital Readmission Reduction Program has recognized the importance of dual eligibility and will assess penalties based on a hospital's performance relative to other hospitals treating a similar proportion of dual eligible patients.[45]

2b3.4a. What were the statistical results of the analyses used to select risk factors?

See the **attached excel document** for risk adjustment model results including the specific risk adjusters, coefficients, p-values, and odds ratios. The final model had a C-statistic of 0.707.

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

We conducted the following analysis using patient-level dual eligibility for social risk factor testing:

- We calculated the proportion of patients with dual eligibility in our sample;
- We compared the discharge to community setting rates, unplanned readmission rates following return to the community, death rates following return to community, and successful discharge to community rates for patients with and without dual eligibility;

---

[43] National Quality Forum, *Evaluation of the NQF Trial Period for Risk Adjustment for Social Risk Factors*, Final Report, 2017.

[44] RTI International, *Developer Response for NQF SDS Trial Period – IRF Readmission NQF #2502,* Submitted to the NQF Standing Committee, 2016.

[45] The Centers for Medicare & Medicaid Services. (2017). *New Stratified Methodology Hospital-Level Impact File User Guide: Hospital Readmissions Reduction Program*. Retrieved from: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Downloads/HRRP_StratMethod_ImpctFile_UG.PDF.

- We added indicators for dual eligibility to our risk adjustment model (described in section **2b3**) and estimated the coefficients and odds ratios of these risk adjusters in the model; and
- We assessed the difference in DTC measure scores estimated with and without dual eligibility adjustment for facilities stratified by the proportion of dual eligible patients with full Medicaid benefits (as this was found to be more prevalent and more strongly related to the DTC outcome).

Table 14 shows the number of dual eligible patients with and without full Medicaid benefits. We found that 11.9% of patient stays were dual eligible with full Medicaid benefits and 5.7% were dual eligible without full Medicaid. The rate of successful DTC for dual eligible patient stays with full Medicaid (59.5%) was 6.32 percentage points lower than the rate for patient stays without dual eligibility (65.7%). The rate of successful DTC for dual eligible beneficiaries without full Medicaid was 2.6 percentage points lower than the rate for patient stays without dual eligibility.

Table 14: Successful Discharge to Community Rates by Dual Eligibility Status

Outcome	Not dual eligible (N=482,227)	Full dual eligible (N=69,923)	Dual without full Medicaid (N=33,552)
Discharge to community setting (codes 01, 06, 81, 86)	75.2%	71.3%	75.2%
Unplanned readmission after return to community	12.1%	16.2%	15.6%
Death after return to community	1.0%	1.0%	1.0%
Successful discharge to community	**65.7%**	**59.5%**	**63.1%**

Source: RTI International analysis of Medicare claims data for index stays in IRF FY2016-2017. (Reference: IB_NQF03).

Table 15 shows the effects of indicators for dual eligibility with full Medicaid and dual eligibility without full Medicaid in our logistic regression model. We found that both dual eligibility indicators were strongly associated with lower odds of successful DTC; as expected, dual eligible with full Medicaid benefits had a stronger negative effect on the DTC outcome. In addition, not shown in **Table 15**, we found that controlling for these two dual eligible indicators increased the model C-statistic by only 0.001.

Table 15: Effect of Dual Eligibility Status Indicators in the Logistic Regression Model

Dual Eligibility Indicator	Estimate	SE	p- value	OR	LCL	UCL
Dual eligible with full Medicaid	-0.2098	0.0095	<.0001	0.811	0.796	0.826
Dual without full Medicaid	-0.0997	0.0130	<.0001	0.905	0.882	0.928

Note: SE=Standard error; OR=Odds ratio; LCL=95% lower confidence limit; UCL=95% upper confidence limit. RTI International analysis of Medicare claims data for index stays in IRF FY2016-2017. (Reference: IB_NQF03)

Table 16 shows the distribution of DTC performance measure scores after adjusting for both dual eligibility indicators and without adjusting for these indicators. The difference between the two sets of scores ranged from -0.87 percentage points to 3.57 percentage points; the mean difference between the two sets of scores was 0.00 percentage points.

Table 16: Distribution of DTC-PAC IRF performance measure scores with and without adjustment for dual eligibility status

DTC Performance Measure Scores	N	Mean	SD	Min	25th Pct	50th Pct	75th Pct	Max
Adjusting for dual eligibility	1,157	64.75	5.34	44.08	61.27	65.20	68.39	82.57
Not adjusting for dual eligibility	1,157	64.74	5.36	43.53	61.29	65.24	68.41	83.35
Percentage point difference in dual-adjusted and non-dual adjusted scores**	1,157	0.00	0.36	-0.87	-0.19	-0.07	0.08	3.57

* Pct = percentile. **Calculated as dual-adjusted score minus non-dual adjusted score for each facility.

RTI International analysis of Medicare claims data for index stays in IRF FY2016-2017. (Reference: IB_NQF03).

We tested the impact of adjustment for dual eligibility on providers stratified by proportion of dual eligible patients with full Medicaid. We stratified providers by dual with full Medicaid for this impact analysis since it was more prevalent and more strongly related to DTC outcomes. We hypothesized that providers with a greater proportion of stays that were dual eligible with full Medicaid would have the greatest improvement in DTC scores after dual adjustment, while providers with a smaller proportion of stays that were dual with full Medicaid may show a decrease in DTC scores after dual adjustment to account for the shift in score distribution compared with the average.

Table 17 shows the differences in DTC measure scores across facilities with and without adjustment for dual eligibility, stratified by the proportion of dual patients with full Medicaid. This table shows that DTC scores of providers with the highest percentages of dual eligible with full Medicaid stays would increase if indicators for dual eligibility were added to our risk adjustment model. For example, DTC scores of providers with more than 25% dual eligible with full Medicaid stays would increase by 0.77 percentage points. As hypothesized, providers with smaller proportions of dual with full Medicaid stays had a decrease in DTC scores after dual adjustment. **Table 17** also shows that as the proportion of stays that are dual with full Medicaid increases, facilities have relatively higher DTC rates when adjusting for dual status compared to not adjusting for dual status.

Table 17: Effect of Adjusting for Dual Eligibility Status on DTC Measure Scores

Facility-level Proportion of Dual Eligible with Full Medicaid Stays	Number of facilities	Mean difference in DTC Scores (DTC score rate adjusted for dual minus score not adjusted for dual) *
≤ 5%	129	-0.26 (0.12)
5%-10%	416	-0.17 (0.09)
10%-15%	281	-0.04 (0.09)
15%-25%	220	0.15 (0.13)
> 25%	111	0.77 (0.66)

Note: *Standard deviation is in parentheses. RTI International analysis of Medicare claims data for index stays in IRF FY2016-2017. (Reference: IB_NQF03).

While dual eligibility had an impact on patient-level DTC outcomes, the impact of adjusting for dual eligibility on facility scores was small. The largest mean improvement in facility DTC scores after dual adjustment was only 0.77 percentage points while the mean score is about 64.82%. The maximum increase in DTC rate with dual adjustment was only 3.57 percentage points, while the mean difference was 0.00 percentage points. Further, we believe adjusting for dual eligibility or other social risk factors for this measure may mask potential disparities in quality of care for dual eligible or other vulnerable patient groups. Therefore, we do not adjust for dual eligibility in our risk adjustment model for the DTC-PAC IRF measure.

2b3.5. 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*)

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

If stratified, skip to 2b3.9

To test the adequacy of this model, we conducted several analyses and tests:

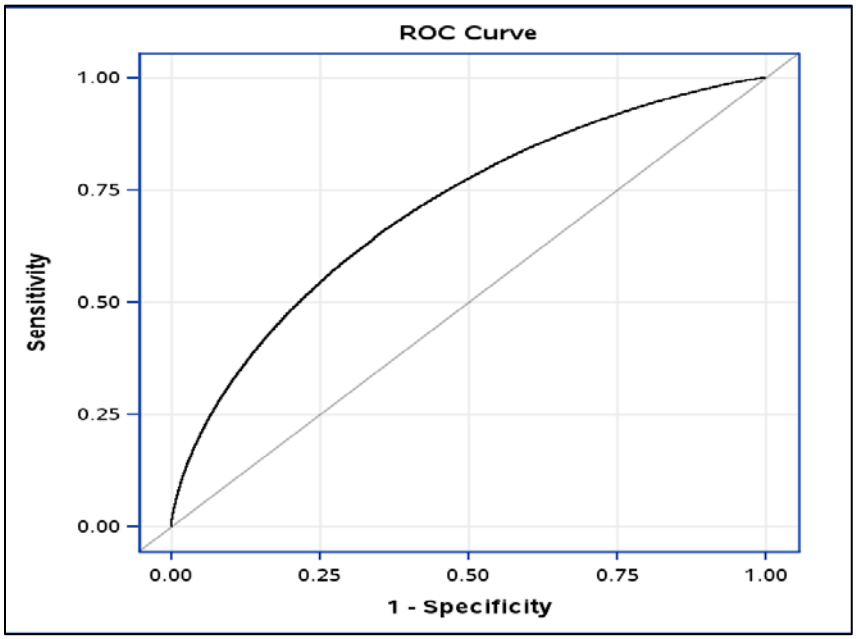
1. **Score Distribution:** We examined the distribution of facility-level observed and risk-standardized discharge to community measure scores.
2. **Model discrimination:** We calculated the C-statistic for the model and generated a receiver operating characteristic curve (ROC). The C-statistic is a measure of how accurately a statistical model can

distinguish between a patient with and without a given outcome. For binary outcomes, the C-statistic is identical to the area under the ROC curve. A C-statistic of 0.50 indicates random prediction, implying the model predicts no better than chance. A C-statistic of 1.0 indicates perfect prediction, implying the model is perfectly predictive of the outcome.

3. **Risk-decile testing and plots:** We calculated the patient-level predicted DTC rates from the logistic model, and assessed the difference between, and ratio of the patient-level observed and predicted DTC rates by deciles of DTC risk. The purpose of this analysis was to assess whether the model worked well across the full range of patients from low to high likelihood of DTC, and to assess for possible over- or under-estimation over the range of risk.

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

The model C-statistic was 0.707, indicating that the model had good predictive ability. **Figure 2** depicts the ROC curve for the model.

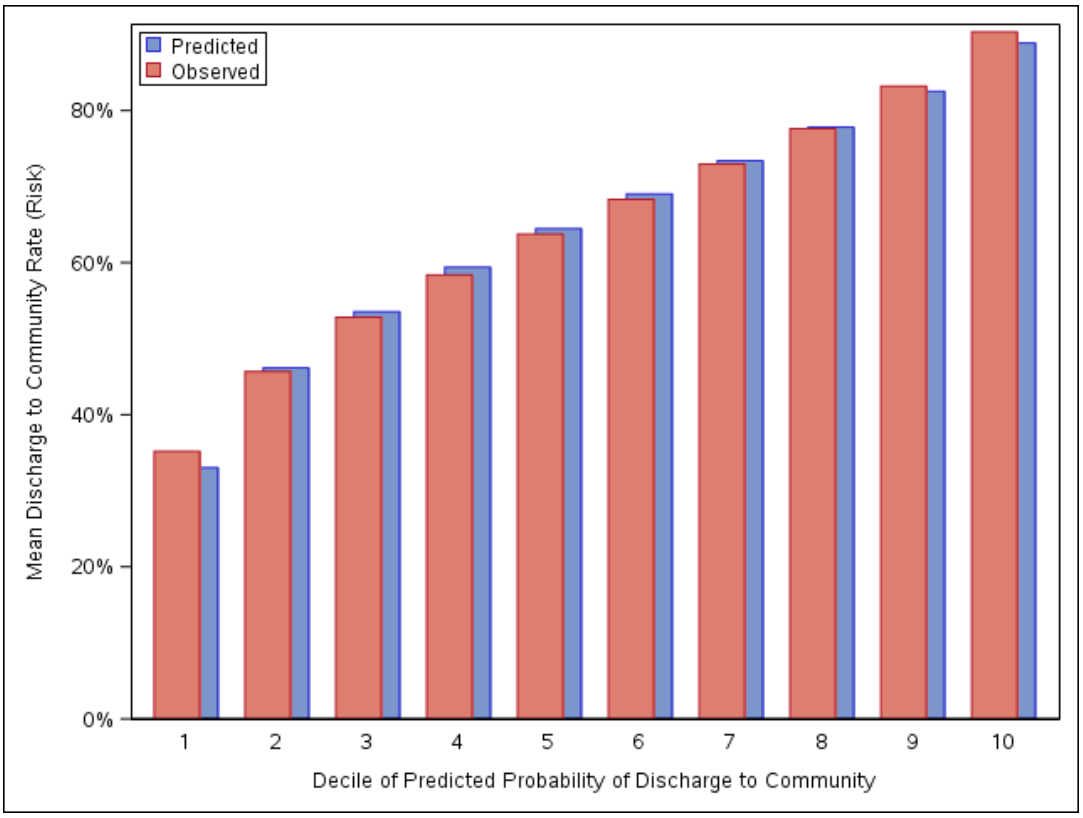


RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: IBNQF10)

Figure 2. Receiver Operating Curve (ROC) for the Patient-Level Logistic Regression Model for the DTC-PAC IRF Measure

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

The patient-level predicted DTC probabilities from the logistic model ranged from 2.95% to 95.81%, , indicating that the model provides a wide range of predictions from low to high. Comparisons of observed and predicted DTC rates by deciles of predicted probability of discharge to community (risk) are displayed in **Figure 3** and **Table 18.** In each risk decile, the patient-level observed and predicted DTC rates are close, with a difference of 1 percentage point or less for eight deciles; the largest difference between observed and predicted rates was 2.2 percentage points. The ratio of observed to predicted rates is close to 1 across risk deciles, with the smallest ratio being .94 in the lowest risk decile. This analysis uses the same approach as the Hosmer-Lemeshow goodness of fit test, without the generation of the Hosmer-Lemeshow statistic. The Hosmer-Lemeshow statistic is not appropriate for very large sample sizes such as the patient sample used in this measure, as even minor deviations can attain statistical significance. Our results demonstrate that this model demonstrates a good predictive ability to distinguish high-risk from low-risk subjects, thus demonstrating good calibration across the range of patients without evidence of concerning under- or over-estimation.



RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: IBNQF10)

Figure 3: Comparison of Patient-Level Observed and Predicted Discharge to Community Rates by Deciles of Predicted Probability of Discharge to Community

Table 18: Comparison of Patient-Level Observed and Predicted Discharge to Community Rates by Deciles of Predicted Probability of Discharge to Community

Deciles of Predicted Probability of Discharge to Community	Number of IRF Discharges	Observed DTC rate	Predicted DTC Rate	Predicted minus Observed DTC Rate (% points)	Predicted DTC Rate/Observed DTC Rate
1	58,570	35.2%	33.0%	-2.2	0.94
2	58,570	45.7%	46.2%	0.5	1.01
3	58,570	52.8%	53.5%	0.7	1.01
4	58,571	58.4%	59.4%	1.0	1.02
5	58,570	63.7%	64.5%	0.7	1.01
6	58,570	68.3%	69.0%	0.7	1.01
7	58,571	73.0%	73.4%	0.4	1.01
8	58,570	77.6%	77.8%	0.2	1.00
9	58,570	83.2%	82.5%	-0.7	0.99
10	58,570	90.3%	88.9%	-1.5	0.98

RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: IBNQF10).

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

See section **2b3.7** for risk decile plots (**Figure 3**) and analysis (**Table 18)**.

2b3.9. Results of Risk Stratification Analysis:

Not Applicable

2b3.10. 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*)

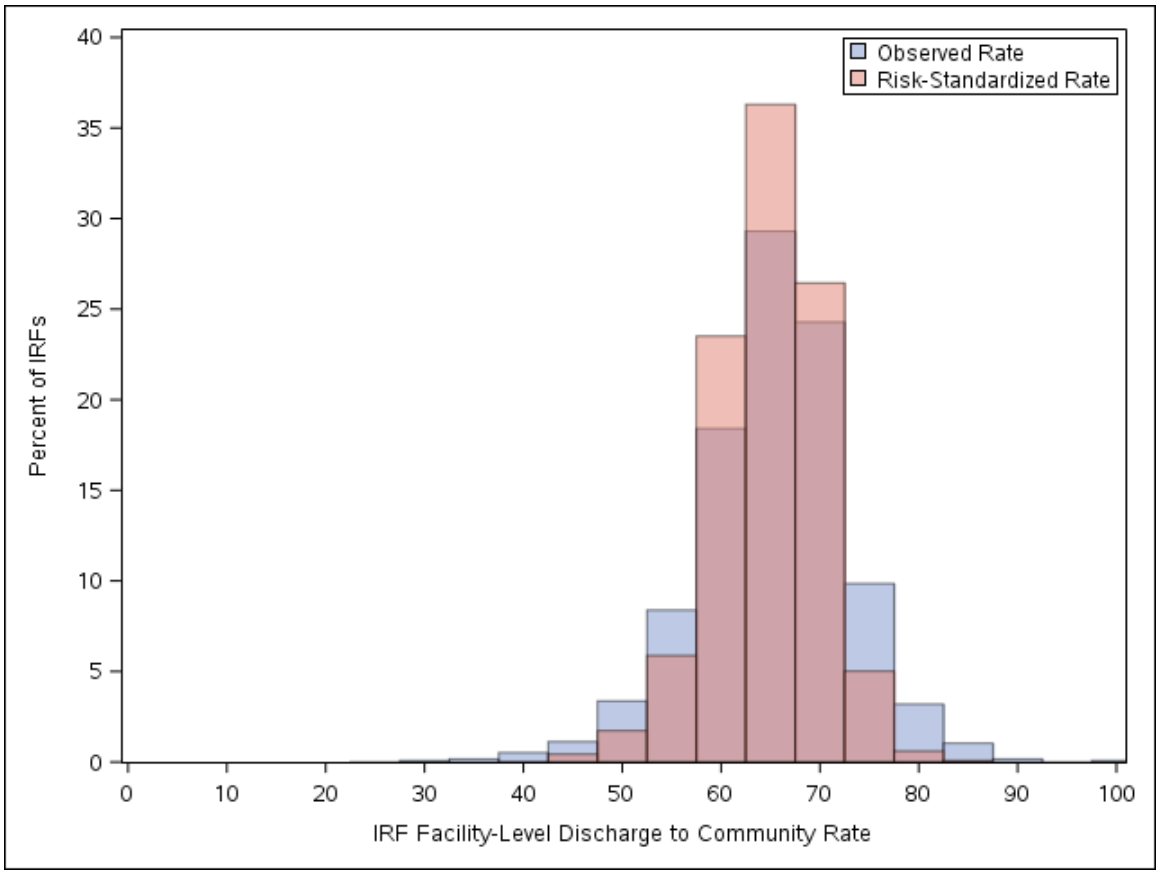
The model discrimination and calibration results demonstrate good model fit and predictive ability across the full range of patients, from low to high risk of successful discharge to community. There was no evidence of extreme under- or over-estimation at the extremes of patient risk.

The distribution of facility-level observed and risk-standardized DTC rates is shown in **Table 19** and **Figure 4.** The observed DTC rate ranges from 29.90% to 100%, with a median of 65.74% and interquartile range of 9.21 percentage points. In contrast, the risk-standardized DTC rate has a much narrower range, from 43.53% to 83.35%, with a median of 65.24% and tighter interquartile range of 7.12 percentage points. The risk-standardized rate also has a smaller standard deviation compared with the observed rate. The narrower distribution of the risk-standardized rates compared with the observed rates is expected based on our risk adjustment methodology with shrinkage towards the mean for low-volume providers.

Table 19: Distribution of Facility-Level Observed and Risk-Standardized Discharge to Community Rates

DTC Rate	N	Mean	SD	Min	25th Pct	Median	75th Pct	Max
Observed	1,157	65.28	7.60	29.90	60.91	65.74	70.12	100.00
Risk-standardized	1,157	64.74	5.36	43.53	61.29	65.24	68.41	83.35

* Pct = percentile. RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: IBNQF10).



RTI Analysis of Medicare Claims File for IRF FY 2016-2017 (reference: IBNQF10).

Figure 4. Distribution of Facility-Level Observed and Risk-Standardized Discharge to Community Rates

2b3.11. 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 that were assessed*)

Not Applicable.

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. 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 used bootstrapping to assess the ability of the DTC measure scores to identify statistically significant differences in provider performance. This analysis was restricted to providers with 25 or more stays during FY 2016-2017 to align with the minimum sample size criterion for public reporting of the measure. Our bootstrapping approach aligns with that used in the IRF all-cause readmission measure (NQF #2502) and other PAC readmission measures. The full set of facilities was sampled with replacement 1,000 times, each facility being randomly selected, or not, for each sample. A range of DTC estimates for each facility was produced varying across samples. Bootstrapping yielded a 95% CI for the providers' risk-standardized DTC estimates. We compared each provider's 95% CI to the national stay-level observed DTC rate (64.82%) to determine if the provider's performance was significantly different from the national rate. Providers whose 95% CI was entirely below the national rate were considered to be significantly worse than the national rate; providers whose CI was entirely above the national rate were significantly better than the national rate; and providers whose CI overlapped the national rate were no different than the national rate.

We calculated the proportion of providers who were significantly different from the national rate.

2b4.2. 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*)

Overall, 95.1% (n = 1,088) of IRFs had performance scores that were significantly different from the national rate, with 44.9% (n = 514) being significantly worse and 50.2% (n = 574) being significantly better than the national rate.

2b4.3. 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 demonstrated excellent ability to identify statistically significant differences in provider performance, providing evidence that the measure can discriminate providers based on quality of care.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note*: This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *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).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with***

more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable, only one set of specifications was used.

2b5.2. 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*)

Not applicable, only one set of specifications was used.

2b5.3. What is your interpretation of the results in terms of the differences in 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*)

Not applicable, only one set of specifications was used.

_____

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

This measure is calculated using Medicare FFS claims data; because submission and completion of claims is tied to hospital reimbursement, missing data are rare. Our measure excludes stays that are missing key measure specification data, under the exclusion criterion of claims data that are problematic. Our data had only 3 stays with a missing CMG code, no stays with unreliable demographic information, and no stays with an undefined CCS grouping (i.e., no invalid claim principal diagnosis codes). These exclusions are not shown individually in **Table 10** because they are grouped under the claims with problematic data exclusion. Thus, missing data are rare and do not have an impact on the measure. Therefore, we did not perform any formal missing data analyses.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; <u>if no empirical sensitivity analysis</u>, identify the approaches for handling missing data that were considered and pros and cons of each*)

See **2b6.1**. Missing data were only present in 3 patient stays, thus do not have an impact on the measure.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; <u>if no empirical analysis</u>, provide rationale for the selected approach for missing data*)

Not Applicable

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in a combination of electronic sources

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

Not applicable. This measure uses Medicare Part A inpatient claims data, which are routinely collected for payment purposes. These data are electronically available from the Centers for Medicare & Medicaid Services (CMS) at no cost beyond that of data processing and can be used to specify, publicly report, and track the measure in a timely fashion. Since data are already collected as part of Medicare's payment process, this measure poses no additional data collection burden on providers. Further, because claims are used for payment, data are complete and subject to audit. In addition to Medicare claims, electronic Medicare enrollment and eligibility data are used.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

None

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
Quality Improvement (Internal to the specific organization)	Public Reporting Inpatient Rehabilitation Facilities Quality Reporting Program https://www.medicare.gov/inpatientrehabilitationfacilitycompare/

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

Name of program and sponsor:
This measure is publicly reported as part of CMS' IRF QRP.
Purpose:
The IRF QRP creates IRF quality reporting requirements, as mandated by Section 3004(a) of the Patient Protection and Affordable Care Act (ACA) of 2010 (H.R. 3590 Health Care Law P.L. Public Law No: 111-148, the Patient Protection and Affordable Care Act). Section 3004(b) of the ACA amended section 1886(j)(7) of the Social Security Act (SSA) requiring the Secretary to establish quality reporting requirements for IRFs. More information about the IRF QRP can be found at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/. More information about ACA Section 3004 (Quality Reporting for Long-Term Care Hospitals (LTCH), Inpatient Rehabilitation Facilities (IRF), and Hospice Programs) can be found at https://burgess.house.gov/uploadedfiles/hr3590_health_care_law_2010.pdf. In addition to tracking quality of care, quality measure data are intended to help consumers make informed decisions when selecting healthcare providers. Most quality measure data from the IRF QRP are publicly reported at https://www.medicare.gov/inpatientrehabilitationfacilitycompare/. IRF quality measure data are also available for download for providers, researchers, and other public at https://data.medicare.gov/data/inpatient-rehabilitation-facility-compare.
Geographic area and number and percentage of accountable entities and patients included:
The IRF QRP includes all IRFs paid under the IRF PPS. Measures are publicly reported for active providers in the reporting period; thus, the number of providers included in the measures can vary by reporting period. Further, the number of providers (and patients) included can vary across published measures because of

differences in measure exclusion criteria and target populations. The DTC-PAC IRF measure was first publicly reported in September 2018 based on 1,157 IRFs, and 585,702 patient stays; of these, data were displayed for active IRFs with 25 or more eligible stays.

Level of measurement and setting:

The DTC-PAC IRF measure is reported at the facility-level for IRF providers.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

Not Applicable

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

Not Applicable

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

Confidential feedback reports on the DTC-PAC IRF measure were provided to all active IRF providers under the IRF PPS in October 2017. Active providers received provider preview reports in June 2018, prior to public reporting of the measure in September 2018. Providers had a 30-day preview period to check their provider preview reports and submit suppression requests if there was evidence of errors in their data. Along with the publicly-reported data, we also include consumer-friendly language to help consumers interpret measure data. Further, we maintain an active provider helpdesk to which providers can submit any questions about the measure, including questions about performance data and interpretation. We provide individual responses to each provider's questions. In addition, CMS conducted an in-person IRF QRP provider training during which providers could ask questions about the measure.[15] Finally, DTC-PAC IRF measure specifications were publicly posted along with the FY 2017 IRF PPS final rule [16] at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/Downloads/Measure-Specifications-for-FY17-IRF-QRP-Final-Rule.pdf. The measure specifications are detailed and precise, allowing stakeholders to replicate measure calculations if they would like.

[15] Centers for Medicare & Medicaid Services Inpatient Rehabilitation Facilities Quality Reporting Program Provider Training: Quality in Post-Acute Care. May 18, 2016. Available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/Downloads/May_2016_IRF_QRP_Provider_Training_Day_One.zip.

[16] Department of Health and Human Services. Centers for Medicare & Medicaid Services. 42 CFR Part 412. Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017; Final Rule. Federal Register, Vol. 81, No. 151.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

See 4a2.1.1.

Confidential feedback reports included the following data: provider number, DTC reporting period start date, DTC reporting period end date, observed number of discharges to community, number of eligible stays, observed discharge to community rate, risk-standardized discharge to community rate, national observed discharge to community rate, comparative performance category, number of IRFs that performed better than

the national rate, number of IRFs that performed no different than the national rate, number of IRFs that performed worse than the national rate, and number of IRFs too small to report.

Provider preview reports included all data included in the confidential feedback reports, along with 95% confidence intervals of the risk-standardized DTC rate.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

In addition to the processes described above, we solicited public comments on the DTC-PAC IRF measure via a 30-day public comment period during November-December 2015, and during the FY 2017 IRF PPS rulemaking process.

4a2.2.2. Summarize the feedback obtained from those being measured.

We received extensive support for implementation of the DTC-PAC IRF measure in the IRF QRP. Comments were received from a range of stakeholders including providers, provider associations, researchers, government agencies, information system vendors, advocacy groups, and individuals. The public comment summary report can be found at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Public-Comment-Summary-Report-for-the-Development-of-a-Discharge-to-Community-Measure.pdf. The FY 2017 IRF PPS final rule with public comments and responses can be found at https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf.[17]

[17] Department of Health and Human Services. Centers for Medicare & Medicaid Services. 42 CFR Part 412. Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017; Final Rule. Federal Register, Vol. 81, No. 151.

4a2.2.3. Summarize the feedback obtained from other users

The public comment summary report can be found at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Public-Comment-Summary-Report-for-the-Development-of-a-Discharge-to-Community-Measure.pdf. The FY 2017 IRF PPS final rule with public comments and responses can be found at https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf.[18]

[18] Department of Health and Human Services. Centers for Medicare & Medicaid Services. 42 CFR Part 412. Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017; Final Rule. Federal Register, Vol. 81, No. 151.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

CMS and RTI International reviewed and considered all public comments before the measure was finalized in the FY 2017 IRF PPS final rule.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

The DTC-PAC IRF measure was recently implemented on October 1, 2016 and publicly reported for the first time in September 2018 using FY 2016-2017 data. Thus, we do not have data to assess trends in performance over time. In the coming years as more data become available, we will examine score distribution and performance improvement over time.

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

No unexpected findings have been noted during implementation of this measure. No unintended impacts on patients have been detected to date.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

Not applicable.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

No

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Not applicable.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment Attachment: NQF_DTC-PAC_IRF_Appendix_RTI.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare & Medicaid Services

Co.2 Point of Contact: Tara, McMullen, Tara.McMullen@cms.hhs.gov, 410-786-8425-

Co.3 Measure Developer if different from Measure Steward: RTI International

Co.4 Point of Contact: Poonam, Pardasaney, pardasaney@rti.org, 312-777-5208-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

The TEP workgroup participated in three meetings, one in-person and two webinars. They provided input on measure conceptualization, definitions, specifications, exclusion criteria, unintended consequences, and other considerations related to development and implementation. The TEP included 17 members, of whom one was a patient representative. The TEP summary report can be found at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Technical-Expert-Panel_-Discharge-to-Community-Quality-Measure.pdf.[19]

TEP members:

1. Susan Adams, PhD, BSN, RN; Vice President of Alliance Integration, Masonicare, Wallingford, CT

2. Greg Arling, PhD; Professor, School of Nursing and Research Associate for the Center for Aging and the Life Course, Purdue University, West Lafayette, IN

3. Dawn Butler, JD, MSW; Director, IU Geriatrics GRACE Training and Resource Center, Adjunct Faculty, Indiana School of Social Work, Indiana University, Indianapolis, IN

4. Michelle Camicia, MSN, CRRN, CCM; Director of Operations, Kaiser Foundation Rehabilitation Center, Nurse Consultant, Vallejo Kaiser Medical Center, Vallejo, CA

5. Susan Hinck, PhD, APRN, GCNS-BC; Director of the Quality Assurance Program, Haven Home Health and Therapy, Ozark, MO

6. Amol Karmarkar, PhD; Assistant Professor, Division of Rehabilitation Sciences, Fellow, Sealy Center on Aging, University of Texas Medical Branch, Galveston, TX

7. Suzanne Kauserud, FACHE, MBA, PT; Vice President of Carolinas Rehabilitation, Surveyor, CARF, Charlotte, NC

8. David Key, DPT; Senior Vice President of Operations for Case Management & Utilization Review, Select Medical Corporation, Mechanicsburg, PA

9. Natalie Leland, PhD, OTR/L, BCG, FAOTA; Assistant Professor, University of Southern California, T.H. Chan Division of Occupational Science and Occupational Therapy, Davis School of Gerontology, Los Angeles, CA

10. Cathy Lipton, MD, CMD; Senior Medical Director, Optum Complex Population Management, Adjunct Clinical Assistant Professor of Medicine, Division of Geriatric Medicine, Emory University School of Medicine, Department of Internal Medicine, Atlanta, GA

11. Rachel Manchester, BSN, MBA, MHA; Regional Director of Home Health Quality, Providence Senior and Community Services, Seattle, WA

12. Keyonna Mayo, BS, Patient representative, Mentor for Women Embracing Abilities Now (W.E.A.N.) and The Dana and Christopher Reeve Foundation, Baltimore, MD

13. Subhadra Nori, MD; Regional Director of the Rehabilitation Medicine Department, Queens Health Network, Elmhurst, NY

14. Terrence O'Malley, MD; Internist/Geriatrician, Massachusetts General Hospital, Boston, MA

15.Lori Popejoy, PhD, APRN, GCNS-BC, FAAN; Associate Professor, Sinclair School of Nursing, University of Missouri, Columbia, MO

16.John Votto, DO, FCCP; President & CEO of Hospital for Special Care, New Britain, CT

17.Christy Whetsell, RN, BSN, MBA, ACM; Director of Case Management, Mid-Atlantic Regional and MountainView Rehabilitation Hospital, President, American Case Management Association, Morgantown, WV

[19] Technical Expert Panel Summary Report: Development of a Discharge to Community Quality Measure for Skilled Nursing Facilities (SNFs), Inpatient Rehabilitation Facilities (IRFs), Long-Term Care Hospitals (LTCHs), and Home Health Agencies (HHAs). RTI International, CMS; February 2016.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released:

Ad.3 Month and Year of most recent revision:

Ad.4 What is your frequency for review/update of this measure?

Ad.5 When is the next scheduled review/update for this measure?

Ad.6 Copyright statement:

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: