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 #: 3477

Corresponding Measures:

De.2. Measure Title: Discharge to Community-Post Acute Care Measure for Home Health Agencies

Co.1.1. Measure Steward: CMS - DCPAC

De.3. Brief Description of Measure: The Discharge to Community-Post Acute Care Measure for Home Health Agencies (DTC-PAC HHA) measure was developed to address the resource use and other measures domain of Discharge to the Community, a domain mandated by the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). The measure was developed using calendar year 2012-2013 data.

This Medicare claims-based outcome measure assesses successful discharge to community from an HHA, with successful discharge to community including no unplanned hospitalizations and no death in the 31 days following discharge. Specifically, this measure reports an HHA's risk-standardized rate of Medicare fee-for-service (FFS) patients who are discharged to the community following an HHA stay, and do not have an unplanned admission 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 based on Medicare FFS claims data and is calculated using two consecutive years of data. This measure submission is based on CY 2015-2016 data; i.e., HHA discharges from January 1, 2015 through December 31, 2016.

The measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the HH Quality Reporting Program finalized in the Calendar Year (CY) 2017 HH Quality Reporting Program (QRP) Final Rule and implementation began October 2016. Confidential feedback reports on measure performance were distributed to HH providers in early 2018. The measure will be publicly reported on the Home Health Compare website (https://www.medicare.gov/homehealthcompare) in January 2019 using CY 2016-2017 data. Four claims-based discharge to community measures were developed for IRF, LTCH, skilled nursing facility, and home health agency settings, respectively 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.

1b.1. Developer Rationale: Successful discharge to community from HHA is widely recognized as an important outcome for residents and their families and an indicator of HH 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. HHAs currently track their discharge to community outcomes for internal quality improvement purposes; however, this tracking commonly assesses initial discharge destination alone and not post-discharge readmissions or death. By assessing whether patients successfully remain in the community for 31 days following discharge, the DTC-PAC HHA 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 HHAs [7-8] (see attached evidence form). The effectiveness of these interventions suggests that improvement in discharge to community rates among HH patients is possible through modifying provider-led processes and interventions. [9-11]

Measuring and publicly reporting the DTC-PAC HHA measure is expected to help differentiate HHAs based on quality of care, and drive improvement in this outcome. The DTC-PAC HHA measure would impact a large number of Medicare beneficiaries and the overall Medicare program. In 2016 Medicare spent $18.1 billion on Medicare fee-for service (FFS) HH care; this care was provided by about 12,200 HHAs nationwide, covering more than 3.4 million beneficiaries.[5] There are no current NQF-endorsed measures assessing successful discharge to community from HHAs.

[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 does not have a simple form for the numerator and denominator—that is, the risk- adjustment method does not make the observed number of community discharges the numerator, and a predicted number the denominator.

The measure numerator is the risk-adjusted predicted estimate of the number of patients who are discharged to the community, do not have an unplanned admission 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 based on Patient Discharge Status Codes 01, 81, the Medicare FFS claim [1]; 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.

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

The numerator uses a model estimated on full national data specific to the PAC setting; it is applied to the HHA's patient stays included in the measure and includes the estimated effect of that HHA. 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 HHA they are discharged from; the effect of the HHA is measured as a positive or negative shift in the intercept term of the equation. The HHA 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 hierarchical logistic regression model presented in this submission are based FY 2016-2017 data.

S.6. Denominator Statement: The target population for the measure is the group of Medicare HH FFS beneficiaries who are discharged from an HHA during the measure time window 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 hierarchical logistic regression model used to calculate the denominator is developed using all non-excluded facility stays in the national data.

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 HH Quality Reporting Program (e.g., excluding HHAs not included in the HHA QRP based on regional location). Stays ending in transfers to the same level of care (i.e., HHA-to-HHA discharge) are excluded, because the HHA episode of care had not ended. We also excluded certain discharge status codes on the HHA FFS claim that indicated that the patient was not appropriate for community discharge (e.g., discharges against medical advice).

Measure exclusion criteria are as follows:

- Age under 18 years;
- 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;
- Stays for patients without continuous Parts A and B FFS Medicare enrollment during the 12 months prior to the HHA admission date and the 31 days after the HHA discharge;
- HHA stays preceded by a short-term acute care or psychiatric stay for non-surgical treatment of cancer;
- Stays ending in transfer to a HHA; and
- Stays with problematic claims data (e.g. anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory).
- Medicare Part A benefits exhausted

De.1. Measure Type: Outcome

S.17. Data Source: Claims, Enrollment Data, Other

S.20. Level of Analysis: Facility

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 home health structures and processes to the likelihood of a patient's successful discharge to the community.
- The developer reviewed additional literature supporting provider structures and processes associated with the discharge to community outcome. The developer states that 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. The developer summarized studies of specific interventions and their impacts as documented in the literature:
    - Adding a dietician can improve nutritional status and reduce the number of hospitalizations within 6-months of the intervention.
    - Cardiac resycnchronization therapy via exercise treatment both in the hospital and at home using telemonitoring. Significant improvement in all domains of quality of living was observed in the home-based group, while the hospital group declared only higher energy levels and less pain.
    - Prescriptions to receive home care were associated with statistical differences in favor of the experimental group for symptoms of delirium, cognitive impairment, and functional status.

- Providing quality primary care in the home setting, nurse practitioners can decrease the number of hospitalizations, 30-day readmissions, and emergency department visits.
- Modified community-based care transitions programs were associated with significantly lower all-cause readmission rate (4%) for those who completed the program than the baseline (15%).
- Home telemonitoring (HTM), which is defined as receiving weekly video televisits, or comprehensive outpatient management (COM), which is defined as patient contact by telephone was associated with increase in QoL and decrease hospital utilization amongst home health patients.
- Home health care for home bound patients with congestive heart failure has shown to to reduce costs and the potential for re-hospitalizations amongst this vulnerable population.
- 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 the Committee satisfied with the summary of the evidence that the developers provided?*
- *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 discharges from HHAs in CY 2015-2016 for HHAs paid under Medicare's HH Prospective Payment System and included in the HH Quality Reporting Program (QRP). All HHAs were included, provided they had eligible stays. Performance scores were calculated for 11,674 HHAs with eligible stays in CY 2015-2016. Facility-level number of HH stays ranged from 1 to 55,067 with a mean of 537.9 and standard deviation (SD) of 1222.6, indicating a skewed sample.
- Observed measure scores ranged from 0% to 100.00%, with a mean of 71.8% and SD of 17.6 percentage points. Risk-standardized performance scores maintained a wide range, 1.7% to 100% with a mean of 77.5% and SD of 16.8 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	0.0%	45.6%	64.4%	71.7%	76.5%	80.1%	83.2%	86.1%	88.8%	92.0%	100%
Risk-Standardized DTC	1.7%	45.3%	65.0%	72.4%	77.0%	80.4%	83.2%	85.8%	88.3%	91.3%	100%

- The developer compared the 95% confidence interval of each provider's performance score to the national stay-level observed DTC rate (77.5%) to determine if the provider's performance was significantly different from the national rate. Overall, 70% of HHAs (n = 7091) had performance scores that were significantly different from the national rate, with 24% (n = 2434) being significantly worse and 46% (n= 4657) being significantly 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 socioecpnomic 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 and race.
    - The successful DTC rate was 74.3% for duals with full Medicaid, 73.9% for duals without full Medicaid, and 80.0 % for non-duals.
    - African-Americans had the lowest DTC rates and Asians had the highest rates, amongst the racial groups with substantial sample size.
    - There was limited variation in DTC rates based on beneficiary residence location or AHRQ SES Index.

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."*

- No additional comments; not aware of new studies
- Almost all of the evidence cited is tangential to the measure; most of the home health-specific studies were examining the effectiveness of home health itself as the intervention; they shed no light on what processes or structures in the delivery of home health may or may not be associated with outcomes. Other studies were based in IRF or SNF settings, so are only suggestive of what may apply to home health. Overall, I count only 2 studies that directly show a link between home health processes and relevant outcomes. This is pretty weak evidence.
- This OUTCOME measure describes the risk-standardized rate of Medicare fee-for-service (FFS) patients/residents/persons who are discharged to the community, and do not have an unplanned (re)admission to an acute care hospital or LTCH in the 31 days following discharge to community, and

remain alive during the 31 days following discharge to community. There is no current NQF-endorsed measures assessing successful discharge to community from HHAs. Domain mandated by IMPACT. (also for IRF, LTCH, SNF). The measure submission is based on CY 2015-2016 data. It was adopted by CMS for the HH Quality Reporting Program in CY17 and implementation began October 2016. Feedback reports on measure performance were distributed to HH providers in early 2018. The measure is, as on Jan 2019, publicly reported on the Home Health Compare website (https://www.medicare.gov/homehealthcompare) using CY 2016-2017 data. The developer provided a logic model linking key home health structures and processes to the likelihood of a patient's successful discharge to the community. Figure 1 (page 28/88; or p 3 of Evidence attachment). Empirical evidence provided came from HH-specific literature, as well as literature from other post-acute care (PAC) and hospital settings, as evidence related to healthcare structures and processes for improving discharge to community outcomes is largely applicable across PAC and hospital settings. Consistent evidence in the literature across PAC settings that rehabilitation interventions, discharge planning, and care coordination can improve discharge to community rates. A review of home health interventions reveals that many factors determine whether or not the patient is successfully discharged to the community. Comprehensive care may include usual care for the specific illnesses, as well as nutritional care, mental health care, medication adherence counselling, telehealth, and fall prevention. The more comprehensive model of care better prepares the patient mentally and physically for discharge helping ensure that they have the necessary knowledge and skills to remain in the community. For Discussion

- Use of studies that are not reflective of Medicare reimbursable home health services including dietitian, nurse practitioners, telephonic phone management, home telemonitoring.
- No evidence that home care prevents death, but can refer to hospice (+intervention)
- No linkage to physician impact as medical oversight required for home care
- Appreciate definitions for unplanned hospitalizations being consistent across PAC; any challenges with this? Suggest more evidence to specifically link Medicare covered serviced to link to staying in the community safely and a discussion on whether the goal really is to prevent death OR should it be a QoL focus Rate as PASS – Relationship between the measured health outcome and at least one healthcare action is demonstrated by empirical data.

- Measure is clearly described. Rationale makes sense. Numerator/Denominator statements are appropriate. Exclusions are reasonable. Numerator description is a bit confusing but after several reads, I understand. Evidence summary is reasonable. Based on the evidence provided, modified community based transition programs that would include primary care in the home setting would positively impact this measure result. PASS

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?)*

- Yes, there is a gap in care as this current measurement does not exist and would be useful
- Yes, there is clearly a performance gap. Some disparities in care are apparently, mainly for dual-eligibility status and possibly by race/ethnicity.
- The developer compared the 95% confidence interval of each provider's performance score to the national stay-level observed DTC rate (77.5%) to determine if the provider's performance was significantly different from the national rate. Overall, 70% of HHAs (n = 7091) had performance scores that were significantly different from the national rate, with 24% (n = 2434) being significantly worse and 46% (n= 4657) being significantly better than the national rate. Disparities The measure does not adjust for social risk factors to avoid masking 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 and race. o The successful DTC rate was 74.3% for duals with full Medicaid, 73.9% for duals without full Medicaid, and 80.0 % for non-duals. o African-Americans had the lowest DTC rates and Asians

had the highest rates, amongst the racial groups with substantial sample size. o There was limited variation in DTC rates based on beneficiary residence location or AHRQ SES. Opportunity for Improvement – MOD/HIGH

- Gaps in care and disparities were clearly defined and evident. The opportunity for improvement is high

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.

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:
  - The materials submitted 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 it is acceptable.

Validity

- Data element:
    - Data element testing was done comparing two-gold standard data element authoritative sources for the discharge to community setting variable using home health claims and OASIS assessment which showed a high rate of agreement.
    - The developers also performed known group testing which showed agreement in the direction of the relationship between specific patient characteristics and the performance rates of the measure.
- 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 five 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 Scientific Methods Panel is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and 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 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: 3477

Measure Title: Discharge to Community-Post Acute Care Measure for Home Health Agencies (HHA)

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:** Medicare eligibility data

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.*

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

   Panel Member #1: The numerator and denominator statements are confusing as written and not complete e.g., no age range noted. The denominator is a bit more defined as to the target population and where it is discharged from. After reading it several times, it appears the denominator is the risk-adjusted expected number of discharges to community. In the denominator the risk adjustment is for patient characteristics (e.g., facility effect is excluded). While the numerator is the risk-adjusted predicted estimate of the patients discharged to community, with no unplanned admissions to acute care hospital or LTCH within a 31-day post discharge and is still alive during the post discharge window. Given the descriptions cited, I have some concerns about consistently implementing this measure as the description of the numerator and denominator are confusing and complicated. The developer needs to be clearer on definitions and they need to be complete.

   Panel Member #2: None

   Panel Member #3: None.

   Panel Member #4: 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

Panel Member #1: For Critical Data Element reliability, developer uses Medicare FFS that are audited. This is appropriate. Measure Score reliability: Developer used signal- ratio and Split-sample testing using ICC to measure reliability. This is appropriate.
Panel Member #2:

- Data Element:
    - 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

Panel Member #3: 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.
Panel Member #4:
Data Element testing
Split-sample ICC
Signal to -noise

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

Submission document: Testing attachment, section 2a2.3

Panel Member #1: Used HHA with 20 or more patient stays in a two-year period. The results are accurate.

Panel Member #2:

- 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 – ICC of 0.91 is strong evidence of 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.

Panel Member #3: Results indicate good precision able to identify systematic differences in quality of care across the HHAs. Reliability testing results indicated good to excellent performance measure score reliability. The mean signal-to-noise ratio was strong in the overall sample, with the signal being

9.7 times as strong as the noise, on average. In the lowest quartile of HHAs the ratio is lower, with the signal bring 2.9 times as strong as the noise, which is still precise. The ICC for the overall HH sample was 0.91, indicating excellent reliability. When examined by facility sample size, the smallest quartile had good reliability (0.86) and the second through fourth quartiles had excellent reliability (0.91-0.96).

Panel Member #4:

Data elements well-matched

Signal to noise median 7.7 overall, by quartile of volume 2.7-5.8-10-17.7

ICC overall 0.91, by quartile of volume 0.86, 0.91, 0.93, 0.96

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 all testing results):

☒**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 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.**

Panel Member #1: Overall signal-to-noise ratio was 9.7 indicating strong signal over noise. ICC between the split-half samples was 0.91 indicating excellent correlation.

Panel Member #2: ICC in split-sample testing was 0.91 which suggests almost perfect measure reliability based on Landis scale.

Panel Member #3: Reliability testing results indicated good to excellent performance measure score reliability. The mean signal-to-noise ratio was strong in the overall sample, with the signal being 9.7 times as strong as the noise, on average. In the lowest quartile of HHAs the ratio is lower, with the signal bring 2.9 times as strong as the noise, which is still precise. The ICC for the overall HH sample was 0.91, indicating excellent reliability. When examined by facility sample size, the smallest quartile had good reliability (0.86) and the second through fourth quartiles had excellent reliability (0.91-0.96).

Panel Member #4: Highly reliable

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

Submission document: Testing attachment, section 2b2.

Panel Member #1: No concerns.

Panel Member #2: none

Panel Member #3: The exclusions are well rationalized and documented with literature and appear to be reasonable.

Panel Member #4: none

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

Submission document: Testing attachment, section 2b4.

Panel Member #1: No concerns.

Panel Member #2: 70% facilities were identified as quality outliers, suggesting that measure can identify meaningful differences in performance

Panel Member #3: none

Panel Member #4: none

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.

Panel Member #1: na

Panel Member #2: NA

Panel Member #3: None.

Panel Member #4: n/a

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

Submission document: Testing attachment, section 2b6.

Panel Member #1: Frequency of missing data is rare 0.2%. The incomplete data are excluded, which is appropriate.

Panel Member #2: Missing data fraction is negligibly small.

Panel Member #4: None.

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

Panel Member #2: 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**

Panel Member #1: No concerns.

Panel Member #2: The testing for the impact of the ONE SES variable tested, dual eligible status, indicated a significant disparity in rates for duals and non-duals, rates of 74.3% vs. 80.0% AFTER adjusting for all other risk factors. The coefficient on the full dual indicator was -0.1489 with small SE (0.003) and p-value <.0001 and OR of 0.862, indicating a nearly 15% 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 over 300 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).

Panel Member #3:

- Hierarchical GLM
- Created non-parsimonious risk adj model, which is appropriate for a risk adj model to ensure adequate case mix adjustment
- Quality is quantified using PE ratio
- Model includes following variables from preceding inpatient stay – age, sex, ADL, LOS for prior hospitalization, CCS diagnosis and procedure categories, CMS HCC categories, #prior acute discharges, #prior ED visits, # SNF visits, # LTCH visits, # dialysis visit
- 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 0.73 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

Panel Member #4: 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 agencies 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.

VALIDITY: TESTING

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

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

☒ **Face validity**

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

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

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

Submission document: Testing attachment, section 2b2.2

Panel Member #1: Used a variety methods (correctness of discharge, known-groups validity, and correlation between established quality and the DTC Expected rate).

Panel Member #2: Validity of elements data assessed by comparing key data elements in database to authoritative source – this was done for outcome variable

Face validity of score assessed using TEP

Assessed convergent validity comparing to other measures

Empiric validity is assessed by assessing predictive validity of risk adjustment model

Panel Member #3: Extensive testing was completed to validate the adequacy of the risk adjustment model, including score distribution, model discrimination and risk-decile testing.

Panel Member #4: Both Face validity and empirical validity (convergent, known-groups) very well done

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

Submission document: Testing attachment, section 2b2.3

Panel Member #1: Acceptable results except for NQF #2505 which was explained in the interpretation discussion.

Panel Member #2:

- 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 (96.6%)
- Convergent validity – showed weak inverse correlation with (1) acute care hospitalization, and (2) ED use; (3) rehospitalization
- Empiric testing of risk adjustment model indicates very good model performance

Panel Member #3: All testing led to similar conclusions and indicated strong validity of the measure.

Panel Member #4: Highly valid

21. **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)

22. **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)

23. **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.)

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

    Panel Member #1: Use of established databases and measures for validity.

    Panel Member #2: Performance of risk adjustment model – which is a measure**.** of predictive validity of measure – is very good

    Panel Member #3: The testing for the impact of the ONE SES variable tested, dual eligible status, indicated a significant disparity in rates for duals and non-duals, rates of 74.3% vs. 80.0% AFTER adjusting for all other risk factors. The coefficient on the full dual indicator was -0.1489 with small SE (0.003) and p-value <.0001 and OR of 0.862, indicating a nearly 15% 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 over 300 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).

    Panel Member #4: I disagree with the handing of dual status but otherwise think this is a very well-developed measure

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

25. **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**

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

ADDITIONAL RECOMMENDATIONS

27. **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.**

Panel Member #1: In my opinion, the measure description and explanation are extraordinarily complex which makes it confusing to those who must implement it. I am wondering whether this measure supplies any information specific to quality improvement initiatives of the organization, or whether it's just a rating as to where an HHA stands as compared to a national average. It seems like it's the latter.

Panel Member #3: Issue here is that even though a clear disparity was found for facilities serving dual eligible members after controlling for facility effect (adjusting for potential poor quality) and other clinical risk factors, because there was no significant impact on overall model results due to removal or addition of ONE risk factor out of over 300 (which is expected!), the decision was made NOT to risk adjust. The potential impact on facilities with high populations of dual or non-duals could very well be very significant given the observed average disparity in rates. How should we deal with this rationale?

Panel Member #4: 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?)*

- No comments
- No concerns about reliability of specifications
- The materials submitted 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. The testing provided was sufficient. Measure score repeatability – (Score-level reliability) was conducted using split sample ICC (below .5 indicates low reliability; the ICC for overall HH samples was 0.91 indicating excellent reliability) Second used a non-standard approach - signal to noise. Results indicated "excellent" ICC's between samples and "good to excellent" mean signal to noise ratios (9.7 times as strong as the noise on average) indicating good precision. Rate – Moderate
- Inferred reliability based on the accuracy of claims data (audited AND used in other NQF claims metrics) is acceptable

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

- No
- No concerns
- No; see comments above
- No. Intraclass Correlation Coefficient is high as well as the SNR

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

- No
- No concerns
- For Discussion:
  - Panel question whether TEP evaluated face validity. In addition, there appears to be two TEP members from HHA (?)

- Face validity was only conducted on the measure concept and results of systematic collection of input was not provided.
- Concern that the risk adjustment model did not include dual status although it was statistically significant in the model Rate for Validity – Moderate

- Valid. The developers rationale for not including Dual status is reasonable

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?"*

- No concerns
- No concerns
- The predicted Discharge to Community rates of the model range from 1.7% to 100%, indicating that this model has a range of predictions and can predict both high and low scores, given the distribution in the observed sample as well. These results regarding distribution and model discrimination indicated that this measure shows meaningful difference in our DTC risk standardized rate for individuals. Of the 10,131 HHAs, most HHA's (70.0%) have scores for this measure that are significantly different (outliers) from the national mean (Table 20). This supports the conclusions that there are meaningful differences in facility-level scores for this measure. Rate as High
- No threats seen

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?"*

- Exclusions are appropriate
- Recommend the committee discuss appropriateness of omitting dual status from risk adjustment.
- Exclusions - For Discussion:
  - Please clarify how the N/D are impacted if patient readmitted DURING the course of home health
  - The second highest frequency of exclusion is for stays where the patient was enrolled in hospice during the post-discharge observation window (5.7%). Many home health patients are chronically complex and may be hospice appropriate. Could referral to hospice be seen as a positive and included favorably in the numerator and denominator?
  - Could Medicare C be a comparison group? Risk Adjustment For Discussion
  - Overall – metrics are complex. •Caregiver ability and willingness AND availability of community resources are critical to staying in the community; how can these be factored in for future?
  - For the specific HHA adjustments, could HHA be compared to a national average.
  - Panel member brought up concern with Dual eligibility; even though a clear disparity was found for facilities serving dual eligible members after controlling for facility effect (adjusting for potential poor quality) and other clinical risk factors, because there was no significant impact on overall model results due to removal or addition of ONE risk factor out of over 300 (which is expected!), the decision was made NOT to risk adjust. The potential impact on facilities with high populations of dual or non-duals could very well be very significant given the observed average disparity in rates.
- Exclusions are approriate. Risk model variables are appropriate.

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 FFS claims from the home health, inpatient, outpatient, and physician office settings 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?"*

- This is feasible as the data is collected already
- Uses readily-available data, although with requirement for 2 years of data this will add burden to calculating measure
- Uses existing systems and data collection already in place, which are electronic. No additional burden to agencies. Rate – High No concerns
- This measure is highly feasible

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

- Home Health 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 HHA measure were provided to all active HH providers under the HH QRP in January, 2018. Active providers received provider preview reports in October, 2018, prior to public reporting of the DTC-PAC HHA measure in January 2019.
    - 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 HHAs that performed better than the national rate, number of HHAs that performed no different than the national rate, number of HHAs that performed worse than the national rate, and number of HHAs too small to report.
- The developer also solicited public comments on the DTC-PAC HHA measure via a 30-day public comment period during November-December 2015, and during the CY 2017 HH QRP 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 can the performance results be used to further the goal of high-quality, efficient healthcare?
- 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

- The DTC-PAC HHA measure was implemented in January 2017 and will be publicly reported for the first time in January 2019 using CY 2015-2016 data. Therefore, there is currently no data to assess trends in performance at this 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?"*

- Pubically reported through the Home Health Quality Reporting Program; public comments have been solicited
- I agree with the methods panelist who brought up concerns about how complex the measure description is (numerator and denominator definitions). This will be a barrier to use, especially for public reporting. How can patients, families or even providers act on the information if they can't understand what it's about. I'd encourage developers to create plain language description of measure to support its transparency and use. Regarding feedback, the report on public comments doesn't show much adjustment of measure design in response to feedback, only rationale for staying the course.
- a1 In 2016, Medicare spent $18.1 billion, to 12,200 HHA covering 3.4 million beneficiaries. This is a growling line of care. Yes, publicly reported as of Jan 2019. This is already implemented measure; just not NQF endorsed. Unclear on use in an accountability program? a2 Yes, feedback reports on the DTC-PAC HHA measure were provided to all active HH providers under the HH QRP in January, 2018. Active providers received provider preview reports in October, 2018, prior to public reporting of the DTC-PAC HHA measure in January 2019. Providers had a 30-day preview period to check these provider preview reports and submit suppression requests if there was evidence of errors in their data. In addition, CMS conducts open door forums during which stakeholders can ask general questions about the measure. CMS also conducted training sessions for home health providers to address the implementation of this measure. Developer also solicited public comment. No summary of feedback from developer provided. For Discussion:
  - Clarification - was once a link to discharge with community resources rated as negative; not consistent with home health; addressed
  - Lessons learned? Other PAC?
- Quite usable and well thought out. Users were involved in development

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."*

- Improvement - similar measures are used in the hospital setting; it would be beneficial for HHAs to have similar data for their organizations as it could show how often patients are discharged prematurely or without appropriate care planning/transitions
- Hard to imagine how useful this measure will be for quality improvement when rates will, by definition, be at least 2 years old. A more timely measure would be more actionable.
- Successful discharge to community is an important outcome for patients; this is a relevant goal. There are issues, such as social determinants that affect the patients ability to safely reside in the home; risk adjustment is addressing most, but not all. Are there discussions to link to value based purchasing. Monitors to prevent agencies for admission-bias? Rating - Pass High; no concerns
- No downsides here

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:
    - 3479: Discharge to Community-Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRF)
    - 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?"*

- Per developer, none have been identified
- No
- This is related to the other discharge to community measures (IRF, LTC, SNF). An item for discussion.
- Not that I can see

Public and Member Comments

NQF received no comments as of: February 1, 2019

Brief Measure Information

NQF #: 3477

Corresponding Measures:

De.2. Measure Title: Discharge to Community-Post Acute Care Measure for Home Health Agencies

Co.1.1. Measure Steward: CMS - DCPAC

De.3. Brief Description of Measure: The Discharge to Community-Post Acute Care Measure for Home Health Agencies (DTC-PAC HHA) measure was developed to address the resource use and other measures domain of Discharge to the Community, a domain mandated by the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). The measure was developed using calendar year 2012-2013 data.

This Medicare claims-based outcome measure assesses successful discharge to community from an HHA, with successful discharge to community including no unplanned hospitalizations and no death in the 31 days following discharge. Specifically, this measure reports an HHA's risk-standardized rate of Medicare fee-for-service (FFS) patients who are discharged to the community following an HHA stay, and do not have an unplanned admission 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 based on Medicare FFS claims data and is calculated using two consecutive years of data. This measure submission is based on CY 2015-2016 data; i.e., HHA discharges from January 1, 2015 through December 31, 2016.

The measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the HH Quality Reporting Program finalized in the Calendar Year (CY) 2017 HH Quality Reporting Program (QRP) Final Rule and implementation began October 2016. Confidential feedback reports on measure performance were distributed to HH providers in early 2018. The measure will be publicly reported on the Home Health Compare website (https://www.medicare.gov/homehealthcompare) in January 2019 using CY 2016-2017 data. Four claims-based discharge to community measures were developed for IRF, LTCH, skilled nursing facility, and home health agency settings, respectively 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.

1b.1. Developer Rationale: Successful discharge to community from HHA is widely recognized as an important outcome for residents and their families and an indicator of HH 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. HHAs currently track their discharge to community outcomes for internal quality improvement purposes; however, this tracking commonly assesses initial discharge destination alone and not post-discharge readmissions or death. By assessing whether patients successfully remain in the community for 31 days following discharge, the DTC-PAC HHA 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 HHAs [7-8] (see attached evidence form). The effectiveness of these interventions suggests that improvement in discharge to community rates among HH patients is possible through modifying provider-led processes and interventions. [9-11]

Measuring and publicly reporting the DTC-PAC HHA measure is expected to help differentiate HHAs based on quality of care, and drive improvement in this outcome. The DTC-PAC HHA measure would impact a large

number of Medicare beneficiaries and the overall Medicare program. In 2016 Medicare spent $18.1 billion on Medicare fee-for service (FFS) HH care; this care was provided by about 12,200 HHAs nationwide, covering more than 3.4 million beneficiaries.[5] There are no current NQF-endorsed measures assessing successful discharge to community from HHAs.

[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 does not have a simple form for the numerator and denominator—that is, the risk- adjustment method does not make the observed number of community discharges the numerator, and a predicted number the denominator.

The measure numerator is the risk-adjusted predicted estimate of the number of patients who are discharged to the community, 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 based on Patient Discharge Status Codes 01, 81, the Medicare FFS claim [1]; 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.

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

The numerator uses a model estimated on full national data specific to the PAC setting; it is applied to the HHA's patient stays included in the measure and includes the estimated effect of that HHA. 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 HHA they are discharged from; the effect of the HHA is measured as a positive or negative shift in the intercept term of the equation. The HHA 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 hierarchical logistic regression model presented in this submission are based FY 2016-2017 data.

S.6. Denominator Statement: The target population for the measure is the group of Medicare HH FFS beneficiaries who are discharged from an HHA during the measure time window 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 hierarchical logistic regression model used to calculate the denominator is developed using all non-excluded facility stays in the national data.

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 HH Quality Reporting Program (e.g., excluding HHAs not included in the HHA QRP based on regional location). Stays ending in transfers to the same level of care (i.e., HHA-to-HHA discharge) are excluded, because the HHA episode of care had not ended. We also excluded certain discharge status codes on the HHA FFS claim that indicated that the patient was not appropriate for community discharge (e.g., discharges against medical advice).

Measure exclusion criteria are as follows:

- Age under 18 years;
- 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;
- Stays for patients without continuous Parts A and B FFS Medicare enrollment during the 12 months prior to the HHA admission date and the 31 days after the HHA discharge;
- HHA stays preceded by a short-term acute care or psychiatric stay for non-surgical treatment of cancer;
- Stays ending in transfer to a HHA; and
- Stays with problematic claims data (e.g. anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory).
- Medicare Part A benefits exhausted

De.1. Measure Type: Outcome

S.17. Data Source: Claims, Enrollment Data, Other

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_evidence_attachment_7.1_DTC_Draft_HH_11.08.18.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 Home Health Agencies (HHAs)

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

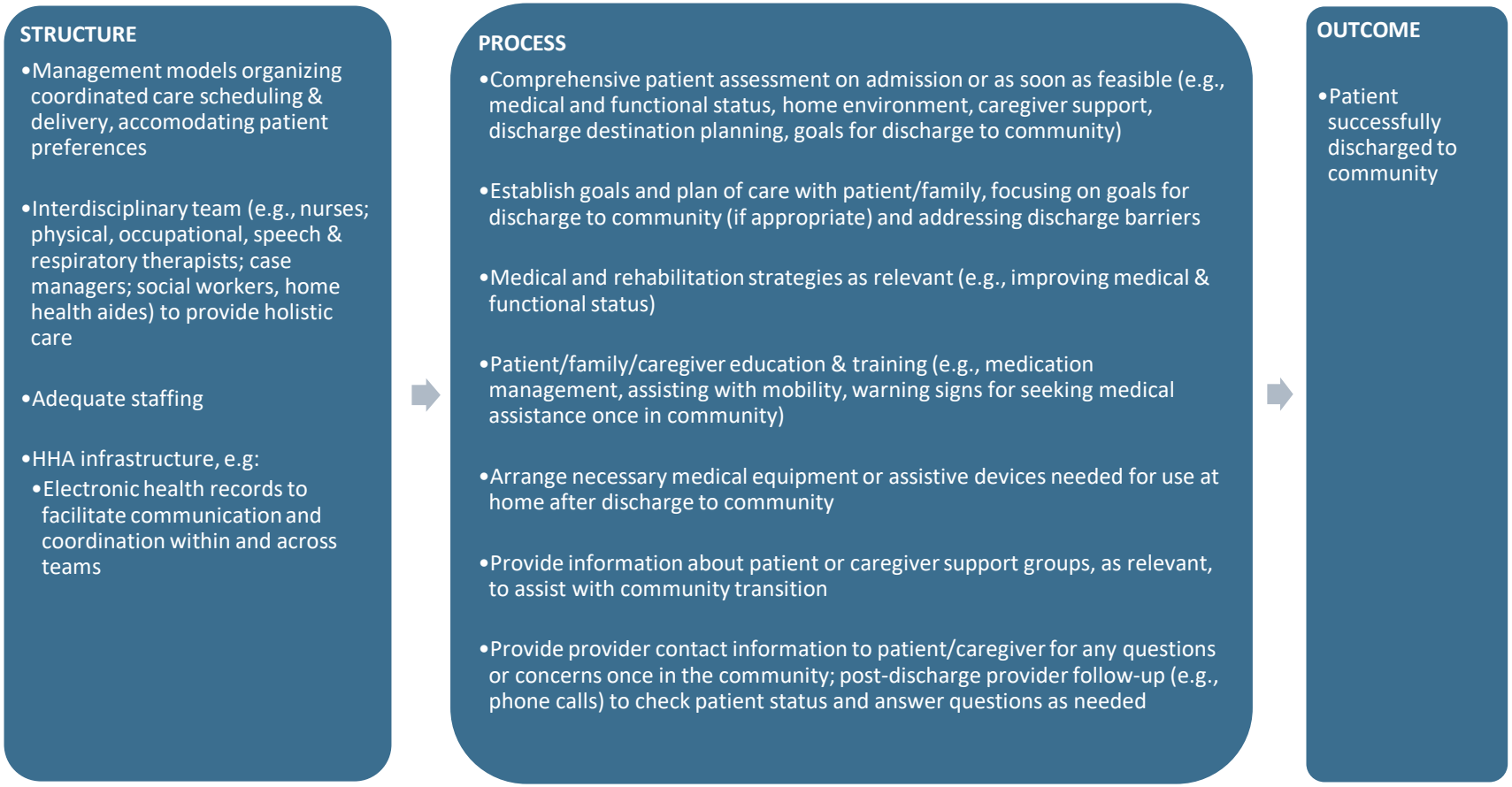
be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

This claims-based outcome measure assesses successful discharge to community from a home health 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 HH settings. Section **1a.2** 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 HHA 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]

[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 HHA 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.)

N/A

****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 HH-specific literature, as well as literature from other post-acute care (PAC) and hospital settings, as evidence related to healthcare structures and processes for improving discharge to community outcomes is largely applicable across PAC and hospital settings. There is consistent evidence in the literature across PAC 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 HHA measure.

A review of home health interventions reveals that many factors determine whether or not the patient is successfully discharged to the community. A comprehensive care approach to home health helps ensure success (i.e., keeping the patient in the community). Comprehensive care may include usual care for the specific illnesses, as well as nutritional care, mental health care, medication adherence counselling, telehealth, and fall prevention. The more comprehensive model of care better prepares the patient mentally and physically for discharge helping ensure that they have the necessary knowledge and skills to remain in the community. [2-8]

Beck et. al [2] tested whether adding a dietician to a discharge team after discharge of geriatric patients improves nutritional status, muscle strength and other patient relevant outcomes. The dietician performed a total of three home visits with the aim of developing and implementing an individual nutritional care plan. In their twelve-week randomized controlled trial of patients seventy and older, researchers found that odds ratio for hospitalization and mortality six months after discharge were 0.367 (0.129; 1.042) and 0.323 (0.060; 1.724) based on receiving additional discharge team support with a dietician versus to discharge team support only. The study concluded that adding a dietician can improve nutritional status and reduce the number of hospitalizations within 6-months of the intervention. [2]

Smolis-Bak et al. [3] assessed the effectiveness of cardiac resycnchronization therapy via exercise treatment. The prospective randomized study was conducted with 52 patients in which the control group (n=26) consisted of patients who had hospital rehabilitation, but no training program after discharge. The treatment group (n=26) underwent initial exercise training in the hospital setting and continued training program at home with telemonitoring 5 times a week for 8weeks. After 3-4months the treatment group achieved better results in VO2 peak, VCO2 peak and treadmill test duration. But after 12months the measurements returned to the baseline values. Other results from the study showed that the effects of hospital and home-based tele-monitoring exercise training found improvements in left ventricular ejection fraction, in both groups (25.3±7.4% to 28.9±9.1%, home-based group, p=0.0213 and 24.9±7.2% to 31.7±10.6%, hospital group, p=0.0001). There was significant improvement in all domains of quality of living was observed in the home-based group, while the hospital group declared only higher energy levels and less pain. The study found that both hospital and home-based programs saw improvements and the home based group experienced additional benefits to quality of life, though the long-term prognosis of patients was not impacted by the kind of intervention they received. [3]

Verloo et al [4] found in a randomized clinical pilot trial using a before/after design was conducted with older patients discharged from hospital who had a medical prescription to receive home care. All patients were monitored for symptoms of delirium using the Confusion Assessment Method. Cognitive and functional

statuses were measured with the Mini-Mental State Examination and the Katz and Lawton Index. A total of 51 patients were randomized into the experimental group (EG) and 52 patients into the control group (CG). Besides usual home care, nursing interventions were offered by a geriatric nurse specialist to the EG. After risk adjustment, statistical differences were found in favor of the EG for symptoms of delirium ($p = 0.046$), cognitive impairment ($p = 0.015$), and functional status ($p = 0.033$), all issues of importance for successfully remaining at home and avoiding hospitalization. This intervention suggests functional and cognitive improvements can be addressed through a nursing intervention. [4]

Echeverry et. al [5] found that by providing quality primary care in the home setting, nurse practitioners can treat homebound patients effectively and decrease the number of hospitalizations, 30-day readmissions, and emergency department visits. Forty patients with a diagnosis of Class III or IV heart failure who were homebound were chosen for this project. The project manager, an Adult-Gerontological nurse practitioner, made home visits to these patients on a monthly and as-needed basis throughout the 3-month project. The rate of hospital admissions, emergency department visits, and 30-day readmissions was reduced by 64%, 85%, and 95%, respectively. Patients were assessed using the Kansas City Cardiomyopathy Questionnaire at the initial visit and at 3 months. The scoring for physical functionality, symptom frequency, and quality of life were improved by 44%, 40%, and 54%, respectively. [5]

In a retrospective analysis, Logue and Drago [6] 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. [6]

Nouryan et. al (2018) in a randomized-controlled study followed 89 Medicare patients with heart failure (HF) after discharge for six months. Patients were randomized to home telemonitoring (HTM) or comprehensive outpatient management (COM). HTM received weekly (video) televisits with daily vital sign monitoring. COM was contacted weekly by telephone. Outcomes included emergency department (ED) and hospital utilization and quality of life (QoL). Thirty-eight percent of HTM had ≥1 ED visit versus 60% of COM ($p = 0.04$), while 48% of HTM had ≥1 hospitalization versus 55% of COM ($p = 0.47$). Length of stay (LOS) (days) was 4.0 for HTM versus 7.4 for COM ($p = 0.39$). Costs were $38,990 for HTM versus $50,943 for COM ($p = 0.91$). QoL improved by -9.66 for HTM and -3.56 for COM ($p = 0.02$). Home telemonitoring is a promising approach to QoL and decrease hospital utilization amongst home health patients. [7]

Punchik et. al [8] completed a retrospective study of healthcare utilization among homebound patients who received home care for CHF from 2012-1015. The outcome measures were number of hospital admissions per month, total number of hospitalization days and days for CHF only, emergency room visits, and overall costs. A comparison was conducted between the 6-month period prior to entry into home care and the time in home care. Over the study period 196 patients were treated by home care for CHF with a mean age of 79.4±9.5 years. 113 (57.7%) were women. Compared to the six months prior to home care, there were statistically significant decreases in hospitalizations (46.3%), in the number of total in-hospital days (28.7%), in the number of in-hospital days for CHF (66.7%), in emergency room visits (47%), and in overall costs (23.9%). Home health

care for home bound patients with CHF can help to reduce costs and the potential for re-hospitalizations amongst this vulnerable population. [8]

The effectiveness of these interventions suggests that improvement in discharge to community rates among HH patients is possible through modified provider-led processes and interventions.

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.

Using an observational study design, Kushner et al.[9,10,11] 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 Uniform Data System for Medical Rehabilitation (UDSMR) data [9,10,11]. 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 [9], 57.8% (of 154) vs. 81.2% (of 151) in stroke patients [10], and 56.9% (of 60) vs. 79.3% (of 58) in geriatric-stroke patients [11]. 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) [9], fewer acute care transfers (27.3% pre-intervention vs. 9.4% post-intervention) [10], reduced length of IRF stay [10,11], and improved Functional Independence Measure (FIM) efficiency [9,10]. 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 [9,10,11], fewer acute care transfers [9,11], fewer long-term care discharges [11], and higher FIM efficiency [9,10,11] compared with case-mix group adjusted national UDSMR data, using a 0.05 significance level.

Berkowitz et al. [12] 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 re-hospitalized 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 re-hospitalization declined 19.4% from 16.5% to 13.3%, and discharges to long-term care fell from 13.8% to 11.5%. [12]

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.

[2] Beck, A., Andersen, U. T., Leedo, E., Jensen, L. L., Martins, K., Quvang, M., Rask, K. O., Vedelspang, A., & Roholt, F. (2015). Does adding a dietician to the liaison team after discharge of geriatric patients improve nutritional outcome: a randomised controlled trial. *Clinical Rehabilitation,* 29(11): 1117–1128.

[3] Smolis-Bak E, Dabrowski R, Piotrowicz E, Chwyczko T., Dobraszkiewicz-Wasilewska B, Kowalik I, Kazimierska B, Jedrzejczyk B, Smolis R, Gepner K, Maciag A, Sterlinski M & Szwed, H. (2015). Hospital-based and telemonitoring guided home-based training programs: effects on exercise tolerance and quality of life in patients with heart failure (NYHA class III) and cardiac resynchronization therapy. A randomized, prospective observation. *International Journal of Cardiology*, 199, 442–447.

[4] Verloo, H., Goulet, C., Morin, D., & von Gunten, A. (2015). Effect Estimation of an Innovative Nursing Intervention to Improve Delirium among Home-Dwelling Older Adults: A Randomized Controlled Pilot Trial. *Dementia and Geriatric Cognitive Disorders Extra*, 5(1): 176–190.

[5] Echeverry, L. M., Lamb, K. V, & Miller, J. (2015). Impact of APN Home Visits in Reducing Healthcare Costs and Improving Function in Homebound Heart Failure. *Home Healthcare Now,* 33(10): 532–537.

[6] 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.

[7] Nouryan CN, Morahan S, Pecinka K, Akerman M, Lesser M, Chaikin D, Castillo S, Zhang M, Pekmezaris (2018) Home Telemonitoring of Community-Dwelling Heart Failure Patients After Home Care Discharge. *Telemedicine and E-Health*

[8] Punchik B, Komarov R, Gavrikov D, Semenov A, Freud T, Kagan E, Goldberg Y, Press Y. *(*2017) Can home care for homebound patients with chronic heart failure reduce hospitalizations and costs? *PLoS One.* 12 (7)

[9] 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.

[10] 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.

[11] 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.

[12] 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.

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 HHA is widely recognized as an important outcome for residents and their families and an indicator of HH 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. HHAs currently track their discharge to community outcomes for internal quality improvement purposes; however, this tracking commonly assesses initial discharge destination alone and not post-discharge readmissions or death. By assessing whether patients successfully remain in the community for 31 days following discharge, the DTC-PAC HHA 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 HHAs [7-8] (see attached evidence form). The effectiveness of these interventions suggests that improvement in discharge to community rates among HH patients is possible through modifying provider-led processes and interventions. [9-11]

Measuring and publicly reporting the DTC-PAC HHA measure is expected to help differentiate HHAs based on quality of care, and drive improvement in this outcome. The DTC-PAC HHA measure would impact a large number of Medicare beneficiaries and the overall Medicare program. In 2016 Medicare spent $18.1 billion on Medicare fee-for service (FFS) HH care; this care was provided by about 12,200 HHAs nationwide, covering more than 3.4 million beneficiaries.[5] There are no current NQF-endorsed measures assessing successful discharge to community from HHAs.

[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 HHA measure is based on national data from CY 2015-2016 Medicare FFS inpatient claims. All HHAs paid under Medicare's HH Prospective Payment System and included in the HH Quality Reporting Program (QRP) are included, provided they had eligible stays. Performance scores were calculated for 11,674 HHAs with eligible stays in CY 2015-2016. Facility-level number of HH stays ranged from 1 to 55,067 with a mean of 537.9 and standard deviation (SD) of 1222.6. Provider characteristics are presented in the testing form (Section 1.5, Table 1). All eligible HH patient stays discharged between January 1, 2015 and December 31, 2016 were included. For patients with multiple HH stays in the measurement period, all eligible stays were included. A total of 6,279,286 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 HHA measure are presented in the testing form (Table 19 and Figure 4). Observed scores ranged from 0% to 100.00%, with a mean of 71.8% and SD of 17.6 percentage points. Risk-standardized performance scores maintained a wide range, 1.7% to 100% with a mean of 77.5% and SD of 16.8 percentage points.

Observed DTC scores by decile were as follows:

Minimum = 0.0%; 10th percentile (pct) = 45.6%; 20th pct = 64.4%; 30th pct = 71.7%; 40th pct = 76.5%; 50th pct = 80.1%; 60th pct = 83.2%; 70th pct = 86.1%; 80th pct 88.8%; 90th pct = 92.0%; max= 100.00%.

Risk-standardized DTC scores by decile were as follows:

Minimum = 1.7%; 10th pct = 45.3%; 20th pct = 65.0%; 30th pct = 72.4%; 40th pct = 77.0%; 50th pct = 80.4%; 60th pct = 83.2%; 70th pct = 85.8%; 80th pct = 88.3%; 90th pct = 91.3%; max = 100%.

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 20 or more stays during CY 2015-2016 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 (77.5%) to determine if the provider's performance was significantly different from the national rate. Overall, 70% of HHAs (n = 7091) had performance scores that were significantly different from the national rate, with 24% (n = 2434) being significantly worse and 46% (n= 4657) being significantly 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 HHA measure has only been implemented in January 2017 and yet to be publicly reported, we do not have data to demonstrate improvement in performance over time. In the coming years when the measure is publicly reported and 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) reported on HHA's population risk-adjusted admission to an acute care hospital, a proxy for an unsuccessful discharge to community, has increased between 2013 and 2016.[5] These results suggest there is room for improvement in managing discharges to limit hospitalizations. The OASIS-based discharge to community measure has also shown an increase in rate of community discharge between 2012 and 2106, from 69.8 to 71.4% [12]. This measure does not assess remaining in community for the length of time as DTC-PAC HH measure but also suggests there is opportunity for continued improvement in rate of successful discharge to community.

[5] March 2018 Report to the Congress: Medicare Payment Policy. Medicare Payment Advisory Commission; 2018.

[12] Abt Analysis of OASIS QMs trend data, CY 2012 – CY 2016.

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 HHA 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 11,674 HHAs and 6,279,286 patient stays using CY 2015-2016 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), 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: CMS Enrollment Database (EDB) (dual eligibility, race, beneficiary 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 26.0% of our sample was dual eligible, with nearly 6.1% having full Medicaid benefits. The majority of our sample was white followed by African-American, and most lived in urban locations.

Some variation in patient-level DTC rates was seen across social risk factor subgroups, with the largest differences seen based on dual status and race. The successful DTC rate was 74.3% for duals with full Medicaid, 73.9% for duals without full Medicaid, and 80.0 % for non-duals. African-Americans had the lowest DTC rates and Asians had the highest rates, amongst the racial groups with substantial sample size. There was limited variation in DTC rates based on beneficiary residence location or AHRQ SES Index, with the first SES index quartile and rural location 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 16). Adjusting for dual status increased the model c-statistic by 0.003 only and had a minimal impact on facility-level performance scores. The difference between dual-adjusted and non-dual—adjusted scores ranged from 0.0 to 3.2 percentage points, with a mean of 0.4 and standard deviation of 0.4 percentage points (testing form, Table 17).

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 whites, Asians have the largest protective effect on the DTC outcome. Non-urban beneficiary residence had a negative impact on the outcome, with rural residence having a larger negative impact than suburban residence. Finally, compared with the highest AHRQ SES index quartile, lower quartiles had a negative effect on the outcome with the lowest SES index group (quartile 1) being the most negative. Adjusting for all social risk factors increased the model c-statistic by 0.0005 only 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 -4.82 to 3.51 percentage points, with a mean of -0.00 and standard deviation of 0.64 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/

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. 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.)*

Measure specifications posted with the CY 2017 HH QRP Final Rule are publicly available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/Downloads/MeasureSpecificationsForCY17-HH-QRP-FR.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_Model_Results_HH_DTC_QM_2018-08-01_FINAL.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.

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 does not have a simple form for the numerator and denominator—that is, the risk- adjustment method does not make the observed number of community discharges the numerator, and a predicted number the denominator.

The measure numerator is the risk-adjusted predicted estimate of the number of patients who are discharged to the community, 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 based on Patient Discharge Status Codes 01, 81, the Medicare FFS claim [1]; 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.

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

The numerator uses a model estimated on full national data specific to the PAC setting; it is applied to the HHA's patient stays included in the measure and includes the estimated effect of that HHA. 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 HHA they are discharged from; the effect of the HHA is measured as a positive or negative shift in the intercept term of the equation. The HHA 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 hierarchical logistic regression model presented in this submission are based FY 2016-2017 data.

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).

Discharge Destination of Community

Discharge to community is determined based on the "Patient Discharge Status Code" from the PAC claim. Discharge to community is defined as discharge to home/self-care [1][2]. Table 1 below lists the Patient Discharge Status Codes used to define community.

Discharge Status Codes Indicating Community Discharge:

01 Discharged to home or self-care (routine discharge)

81 Discharged to home or self-care with a planned acute care hospital readmission

Unplanned Admissions in the 31-Day Post-Discharge Observation Window

A patient who is discharged to the community is not considered to have a successful discharge to community outcome for this measure if they have a subsequent unplanned admission to an acute care hospital or LTCH in the post-discharge observation window, which includes the day of discharge and the 31 days following day of discharge.

We identify unplanned admissions based on the planned readmissions algorithm used in the following post-acute care readmission measures, endorsed by the National Quality Forum (NQF): (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 (NQF #2502); (iii) All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Long-Term Care Hospitals(NQF #2512) and (iv) Rehospitalization During the First 30 Days of Home Health (NQF #2380).[3][4][5][6]

These PAC readmission measures are based on the Hospital-Wide All-Cause Readmission Measure (HWR) (NQF #1789),[7] with some additions made for the SNF, IRF, and LTCH setting measures.[8] The 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 initially developed with ICD-9-CM (International Classification of Diseases, Ninth Revision, Clinical Modification) procedure and diagnosis codes, it was transitioned using the ICD-9-CM to ICD-10-CM cross-walk. All analyses presented in this submission are based on both ICD-9-CM and ICD-10-CM codes.

Death in the 31-Day Post-Discharge Observation Window

A patient who is discharged to the community 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.

Measure Time Window

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

References

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

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

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

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

7- Hospital-Wide All-Cause Readmission Measure (HWR) (CMS/Yale) (NQF #1789). www.qualityforum.org/QPS/1789

8 - 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.

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

The target population for the measure is the group of Medicare HH FFS beneficiaries who are discharged from an HHA during the measure time window 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 hierarchical logistic regression model used to calculate the denominator is developed using all non-excluded facility stays in the national data.

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).

As previously stated, the measure does not have a simple form for the numerator and denominator. 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 a HHA during the measure time window and are not excluded based on the measure exclusion criteria. The target population for the analyses in the submission includes HHA discharges from January 1, 2015 through December 31, 2016 (i.e., CY 2015-2016).

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 HH Quality Reporting Program (e.g., excluding HHAs not included in the HHA QRP based on regional location). Stays ending in transfers to the same level of care (i.e., HHA-to-HHA discharge) are excluded, because the HHA episode of care had not ended. We also excluded certain discharge status codes on the HHA FFS claim that indicated that the patient was not appropriate for community discharge (e.g., discharges against medical advice).

Measure exclusion criteria are as follows:

- Age under 18 years;
- 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;
- Stays for patients without continuous Parts A and B FFS Medicare enrollment during the 12 months prior to the HHA admission date and the 31 days after the HHA discharge;
- HHA stays preceded by a short-term acute care or psychiatric stay for non-surgical treatment of cancer;
- Stays ending in transfer to a HHA; and
- Stays with problematic claims data (e.g. anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory).
- Medicare Part A benefits exhausted

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 HHA measure are listed below, along with the rationale for each exclusion. The measure exclusion criteria are determined by processing Medicare claims and eligibility data to determine whether the individual exclusion criteria are met. All measure exclusion criteria are based on administrative data.

Exclusions for the discharge to community measure are listed below, along with the rationale for each exclusion.

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 post-acute Medicare FFS population.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

2. 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 HHA access for patients discharged from psychiatric hospitals.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

3. Discharges against medical advice

Rationale: Patients who discharge themselves against medical advice are excluded because their 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 admissions or death, depending on their medical condition, or due to potential non-adherence or non-compliance with care recommendations.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

4. Discharges to disaster alternative care sites or federal hospitals

Rationale: Patients discharged 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 necessary inpatient claims for these patients.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

5. 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: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

6. Patients discharged 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 provider, makes the final decision of discharge to hospice-home or hospice-facility.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

7. Patients not continuously enrolled in Parts A and B FFS Medicare (or those enrolled in Part C Medicare Advantage) for the 12 months prior to the post-acute admission date, and at least 31 days after post-acute discharge date

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

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

8. Patients whose prior short-term acute care or psychiatric stay was for non-surgical treatment of cancer

Rationale: Patients whose 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 patients is consistent with the hospital-wide and post-acute readmission measures.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

9. Post-acute stays that end in transfer to the same level of care

Rationale: HHA stays that end in transfer to another HHA are excluded from the measure because the HHA episode has not ended. For a HHA episode that involves transfer to another HHA, only the final HHA provider is included in the measure.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

10. Post-acute 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)

Rationale: This measure requires accurate information from the post-acute stay and prior short-term acute care stay in the elements used for risk adjustment. No-pay post-acute stays involving exhaustion of Part A benefits are also excluded.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

11. Medicare Part A benefits exhausted

Rationale: Patients who have exhausted their Medicare Part A coverage during the PAC stay are excluded because the discharge destination decision may be related to exhaustion of benefits.

Data Source: Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

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

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 HHA measure is risk adjusted. To develop the risk-adjustment model for this measure, we analyzed Medicare home health, inpatient claims, outpatient, and carrier claims. 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 (CY 2015-2016).

The risk model employs the following sets of covariates:

(1) Demographics

(a) Age and sex

(b) Enrollment status

(c) Activities of daily living scores

(2) Care received during the prior proximal hospitalization (if relevant)

(a) Length of prior proximal hospitalization

(b) Clinical classification software (CCS) diagnosis and procedure categories during prior proximal hospitalization

(3) Other care received within one year of the HH stay

(a) Number of prior acute discharges

(b) Number of outpatient emergency department visits

(c) Number of skilled nursing facility visits

(d) Number of long-term care hospital visits

(e) Number of inpatient dialysis sessions

(f) Hierarchical condition categories (HCC) comorbidities

We used a hierarchical logistic regression model to predict the probability of discharge to community. Patient characteristics related to discharge and a marker for the specific discharging HHA 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 HHA. 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 HHA effect, hierarchical modeling accounts for the known predictors of discharge to community, on average, such as patient characteristics, the observed facility rate, and the number of HHA stays eligible for inclusion in the measure. The estimated HHA effect is determined mostly by the HHA's own data if the number of patient discharges is relatively large (as the estimate would be relatively precise) but is adjusted toward the average if the number of patient discharges 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 is a measure of the degree to which discharges to community are higher or lower than what would otherwise be expected. This

standardized risk ratio is then multiplied by the mean discharge to community rate for all HHA stays for the measure, yielding the risk-standardized discharge to community rate for each HHA.

Please note that 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.

Using a two-year reporting period, CMS intends to publicly report the performance of Medicare-certified home health agencies (with at least 20 home health stays) on the DTC-PAC HHA measure under three performance categories: "Better than Expected," "Same as Expected," and "Worse than Expected". Pursuing a categorical reporting method is consistent with the Hospital-Wide All-Cause Readmission Measure.

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

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

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

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

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

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

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

predj = Slogit-1(u + wi + B*Zij) (2)

where the sum is over all stays in HHAj, and wi is the random intercept.

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 HH-specific effect included in the predictions. This produces the expected number of discharges at the average HHA.

To calculate the expected number expj, we used the following equation:

expj = Slogit-1 (u + B*Zij) (3)

Step 5: Calculate the standardized risk ratio for each HHA, as the ratio of the predicted to expected number of discharges to community.

To calculate the HHA-wide standardized risk ratio, SRRj, we used the following equation:

SRRj = predj/expj (4)

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

To aid interpretation, the HHA-wide standardized risk ratio, SRRj, obtained from equation (4) is then multiplied by the overall national raw discharge to community rate for all HH stays, ?, to produce the HHA-wide risk-standardized discharge to community rate (RSRj).

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

RSRj = SRRj*? (5)

NOTE: It is important to clarify that the DTC-PAC HHA measure is specific to HHA 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

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

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

If other, please describe in S.18.

Claims, Enrollment Data, Other

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.

The following are the specific files used for the DTC-PAC HHA measure and links to the documentation:

Medicare Inpatient claims (Standard Analytical Files) for hospital claims, including index HHA 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 following web page includes data dictionaries for the Standard analytical files:

- Home Health RIF: http://www.resdac.org/cms-data/files/hha-rif
- Inpatient RIF: http://www.resdac.org/cms-data/files/ip-rif
- Outpatient RIF: http://www.resdac.org/cms-data/files/op-rif
- Carrier (Physician Office) RIF: http://www.resdac.org/cms- data/files/carrier-rif

Medicare Enrollment Database:

- Information about the Enrollment Database may be found at: https://aspe.hhs.gov/centers-medicare-medicaid-services

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.)*

2. Validity – See attached Measure Testing Submission Form

DTC_NQF_Testing_Attachment_7.1_DTC_HH_2018-08-01_Final.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 Home Health Agencies
Date of Submission:

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 e 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 f 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; g

AND

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

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; i,j 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 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

e. 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).

f. 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.

g. 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.

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

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

j. 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: Home Health Outcome and Assessment Information Set (OASIS) Version: C1; Online Survey Certification & Reporting System (OSCAR); Common Medicare Environment database

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).

This measure is based on Medicare's eligibility database as well as Medicare fee-for service (FFS) administrative claims from home health, inpatient, outpatient, and physician office settings. The eligibility files provide beneficiary-level information such as date of birth, date of death, sex, reasons for Medicare eligibility, enrollment histories in Medicare Parts A and B, and beneficiary dual eligibility status. The Medicare FFS claims files provide information about each home health and PAC stay, including dates of admission and discharge, diagnoses and procedures, and indicators for care received in the intensive care unit, coronary care unit, and emergency department. Furthermore, claims from all three file settings are used to construct for each patient a complete history of care before the index home health stay, which is used for constructing risk adjustment variables. 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 originally developed using calendar year (CY) 2012-2013 data. This measure submission is based on CY 2015-2016 data; i.e., home health discharges from January 1, 2015 through December 31, 2016. We used the data sources listed below to develop the claims analytic file for measure specification and testing.

- Medicare Enrollment Database: Information is available at: http://aspe.hhs.gov/datacncl/datadir/cms.htm
- Medicare Claims data from home health, inpatient, outpatient, and physician office settings. Data dictionaries for the Standard analytical files are provided online by Centers for Medicare & Medicaid Services (CMS) contractor, Research Data Assistance Center (ResDAC), at the University of Minnesota
    - Health Research Identifiable File (RIF): http://www.resdac.org/cms-data/files/hha-rif/data-documentation
    - Inpatient RIF: http://www.resdac.org/cms-data/files/ip-rif/data-documentation
    - Outpatient RIF: http://www.resdac.org/cms-data/files/op-rif/data-documentation
    - Carrier (Physician Office) RIF: http://www.resdac.org/cms-data/files/carrier-rif/data-documentation
- AHRQ CCS groupings of ICD-9 and ICD-10 codes: 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-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:

- Home Health Outcome and Assessment Information Set (OASIS) Version C1: documentation available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/HHQIArchives.html
    - We used OASIS items for measure validation, including OASIS start of care date, reason for assessment, Integumentarey Status (M1306), Activities of Daily Living (M1800) as a proxy for Section GG discharge functional status, and Emergent Care (M2310) CY 2015-2016 data.
- Common Medicare Environment (CME) database: CMS has designated the CME database as the single, enterprise-wide authoritative source for Medicare beneficiary enrollment and demographic data. The CME database integrates and standardizes different types of beneficiary data from CMS legacy systems. The CME database receives information from the EDB and also contains additional information not available in the EDB. We extracted patient-level dual eligibility information from the CME database for the social risk factors analysis. Description of the CME available at:

https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf

- Online Survey Certification & Reporting System (OSCAR): contains data on characteristics of home health agencies, including the name and address of the facility and the type of Medicare services the home health agency provides, among other information. This information was used to generate **Table 1** (provider-level characteristics).

1.3. What are the dates of the data used in testing? Calendar year 2015-2016 (i.e., home health discharges from January 1, 2015 through December 31, 2016. By having a two year data collection period for the HH setting, there is a larger number of eligible home health stays for each home health agency. Therefore, more agencies are able to meet the public reporting threshold (i.e., a minimum of 20 home health stays).

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 home health agencies paid under Medicare's HH Prospective Payment System and included in the HH Quality Reporting Program (QRP) were included, provided they had eligible stays. A total of 11,674 home health agencies (HHAs) with eligible stays in CY 2015-2016 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 11,674 HHAs included in Specification and Testing of the Discharge to Community-PAC HH QRP Measure (CY 2015-2016)

Characteristic	***K* (%)**	**Characteristic**	***K* (%)**
Census Region		**Ownership type**	
New England	389 (3.3%)	Government	538 (4.6%)
Mid Atlantic	522 (4.5%)	Not-For-Profit	1,804 (15.5%)
East North Central	2,377 (20.4%)	Proprietary	9,332 (79.9%)
West North Central	750 (6.4%)		
South Atlantic	1,868 (16.0%)		
East South Central	432 (3.7%)		
West South Central	3,128 (26.8%)	**Rurality**	
Mountain	720 (6.2%)	Rural	2,214 (19.0%)
Pacific	1,436 (12.3%)	Urban	9,460 (81.0%)
US Territories	52 (0.4%)		

Acumen Analysis of Medicare Claims File for HH and Online Survey Certification and Reporting (OSCAR) data, CY 2015-2016.

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

Measure specification and testing was based on national data. All eligible HH patient stays discharged between January 1, 2015 and December 31, 2016 were included in the measure. This was the most recent data time range that was available for measure testing. For patients with multiple HH stays in the measure calculation period, all eligible stays were included. A total of 6,279,286 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 HH 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 **attached excel document.** Throughout this form, we use "*N*" to refer to number of patient stays.

Table 2. Demographic Characteristics of HH Patient Stays Included in Specification and Testing of the Discharge to Community-PAC HH QRP Measure (N = 6,279, 286)

Characteristic	N (%)	Characteristic	N (%)	
Male	2,338,796 (37.2%)	**Race**		
Female	3,940,490 (62.8%)	White	5,156,630 (82.1%)	
Male < 65	345,496 (5.5%)	Black	744,178 (11.9%)	
Female < 65	411,993 (6.6%)	Other	378,478 (6.0%)	

Acumen Analysis of Medicare Claims File for HH, CY 2015-2016.

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 6,279,286 patient stays at the 11,674 Home Health Agencies described above were included in all aspects of measure testing, with differences described below.

1. **Critical data element validity testing:** Some aspects of validity testing (**2b1**) utilized OASIS-C1 assessment data.
    a. **Agreement between Claims Discharge Status Code and Discharge Status on OASIS-C1 Assessment**: We matched index HH stays with OASIS-C1 Assessments based on patient and facility identifiers, and the start of care date. A total of 6,154,523 index HH stays were successfully matched to OASIS-C1 assessments and were included in data element validity testing of the claims discharge status code presented in section **2b1.3** (**Table 5**).
    b. **Known-groups Validity Testing (Section 2b1.3):**
        i. Data for the M2310 (Reason for Emergent Care) item were collected from transfer to inpatient facility or discharge from agency assessments. Thus data element validity testing was conducted on HH stays ending in a transfer or discharge assessment (n = 5,634,182).
        ii. Data on functional status (M1800-M1890) and cognitive function (M1700) items were collected from discharge from agency assessments. Thus data element validity testing was conducted on HH stays ending in a discharge assessment where data was available (n= 4,542,500).
2. **Performance measure score testing:** Facility-level testing (sections **2a2**, **2b1.3**, **2b4.2**) were restricted to 10,131 HHAs with 20 or more stays in CY 2015-2016; we applied this restriction

because this measure will only be publically reported for providers with a minimum of 20 stays in the two-year measure calculation period.

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.

Data for rural location, and Medicaid dual status were readily available through the enrollment database (EDB) and Common Medicare Environment (CME) and analyzed during the measure testing process. While a scoping review[1] found general agreement that persons of lower socioeconomic status are not disadvantaged in terms of HH care services, there is a well-documented socioeconomic gradient seen with primary and acute care services. In addition, due to the availability of data and previous evidence on the importance of dual eligibility in similar claims-based measures, we test patient-level dual eligibility. For example, starting in fiscal year 2019, CMS's Hospital Readmission Reduction Program will assess penalties based on a hospital's performance relative to other hospitals treating a similar proportion of dual eligible patients.

We then conducted the following analysis for each social risk factor:

- Calculated the prevalence of patients with each social risk factor;
- Compared the unplanned admission rates following return to the community and successful discharge to community rates for patients with and without each social risk factor;
- Calculated the distribution of agency-level SDS factor prevalence;
- Calculated the difference in agency-level risk-standardized rates; and
- Assessed the difference in DTC measure scores estimated with and without dual eligibility adjustment for agencies 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).

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. Critical Data Element Reliability

---

[1] Goodridge et al. Socioeconomic disparities in home health care service access and utilization: A scoping review 2012: *International J. Nursing Studies* 49(10); 1310-19

[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

The DTC-PAC HH 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 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 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 inpatient, outpatient, carrier, and skilled nursing facility claims in the year preceding the HH stay, the ability of HH providers to influence the data to improve their measure performance is minimized. The measure outcome is based on the Patient Discharge Status Code on the HH claim indicating discharge to a community setting. However, 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 HH discharge. Presence of such a claim negates the discharge to community identified based on the HH claim. Section **2b1.1** discusses empirical data element validity testing of the Patient Discharge Status Code on the HH 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 in quality measures. These measures include the following: (i) Hospital-Wide All-Cause Readmission Measure (HWR) (CMS/Yale) (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 (NQF #2502); (iv) All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Long Term Care Hospitals (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 a form of measure score repeatability by assessing agreement between a facility's performance measure scores based on randomly-split independent patient stay subsets; and (ii) we examined signal-to-noise ratios. Performance-measure score reliability testing was restricted to the 10,131 HHAs that had 20 or more patient stays in the CY 2015-2016 measurement period, as only HHAs with a minimum of 20 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 score 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).

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

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.

---

3 Hospital-Wide All-Cause Readmission Measure (HWR) (CMS/Yale) (NQF #1789). http://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 (NQF #2502). http://www.qualityforum.org/QPS/2502

6 All-Cause Unplanned Readmission Measure for 30 Days Post Discharge from Long Term Care Hospitals (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

In addition to calculating signal-to-noise for the overall facility sample of 10,131 HHAs, 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.

b. **Split-sample Reliability Testing**: This testing examined agreement between two performance measure scores for a facility calculated using 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) used for measure calculation. We randomly divided each facility's CY 2015-2016 patient stays into halves stratifying by calendar year, thus ensuring that patient stays within each calendar 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 irrespective of sample size. To do this, we first split our sample of 10,131 HHAs 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 the **signal-to-noise ratio** of performance measure scores for the overall sample of 10,131 HHAs included in this testing, and by sample size quartile.

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

Facility Sample	*K*	Mean (SD)	Min	25th Pct*	Median	75th Pct	Max
Overall	10,131	9.7 (7.8)	0.5	4.0	7.7	13.2	77.8
20-94 Stays (quartile 1)	2,520	2.9 (1.5)	0.5	1.8	2.7	3.9	9.7
95 - 243 Stays (quartile 2)	2,546	6.0 (2.5)	0.6	4.3	5.8	7.5	19.9
244-625 Stays (quartile 3)	2,533	10.3 (3.5)	1.5	8.0	10.0	12.3	33.0
626-55,067 Stays (quartile 4)	2,532	19.6 (8.0)	1.1	14.0	17.8	23.4	77.8

Acumen Analysis of Medicare Claims File for HH, CY 2015-2016. Note: The reliability testing was performed among HHAs with >= 20 stays. As a result, 1,543 HHAs were excluded from the analysis due to insufficient sample size.

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 10,131 HHAs included in this testing, and by sample size quartile. The ICC (2,1) and ICC (3,1) estimates were identical for all sample size quartiles. The ICC in the overall sample was 0.91, with a 95% confidence interval (CI) of 0.91 to 0.91. When examined

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

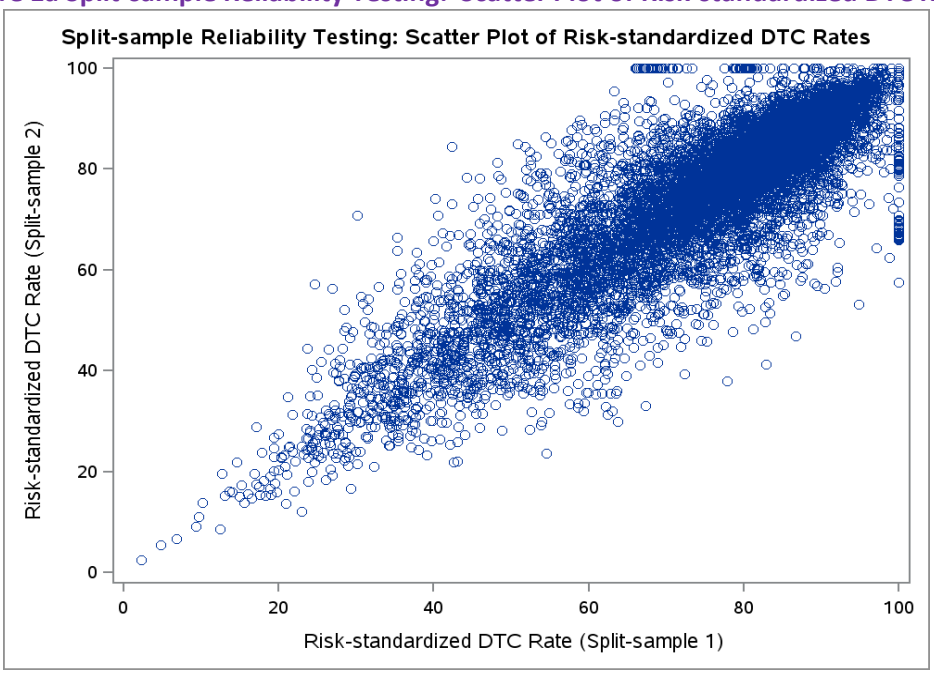
by sample size quartile, the lowest ICC was 0.86 in the first sample size quartile. **Figures 1a and 1b** show scatter plots of the split-half scores overall, and by sample size quartile.

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

Facility Sample	*K*	ICC (2,1) (95% CI)	ICC (3,1) (95% CI)
Overall	10,131	0.91(0.91-0.91)	0.91(0.91-0.91)
20-94 Stays (quartile 1)	2,520	0.86(0.85-0.87)	0.86(0.85-0.87)
95 - 243 Stays (quartile 2)	2,546	0.91(0.90-0.91)	0.91(0.90-0.91)
244-625 Stays (quartile 3)	2,533	0.93(0.92-0.94)	0.93(0.92-0.94)
626-55,067 Stays (quartile 4)	2,532	0.96(0.95-0.96)	0.96(0.95-0.96)

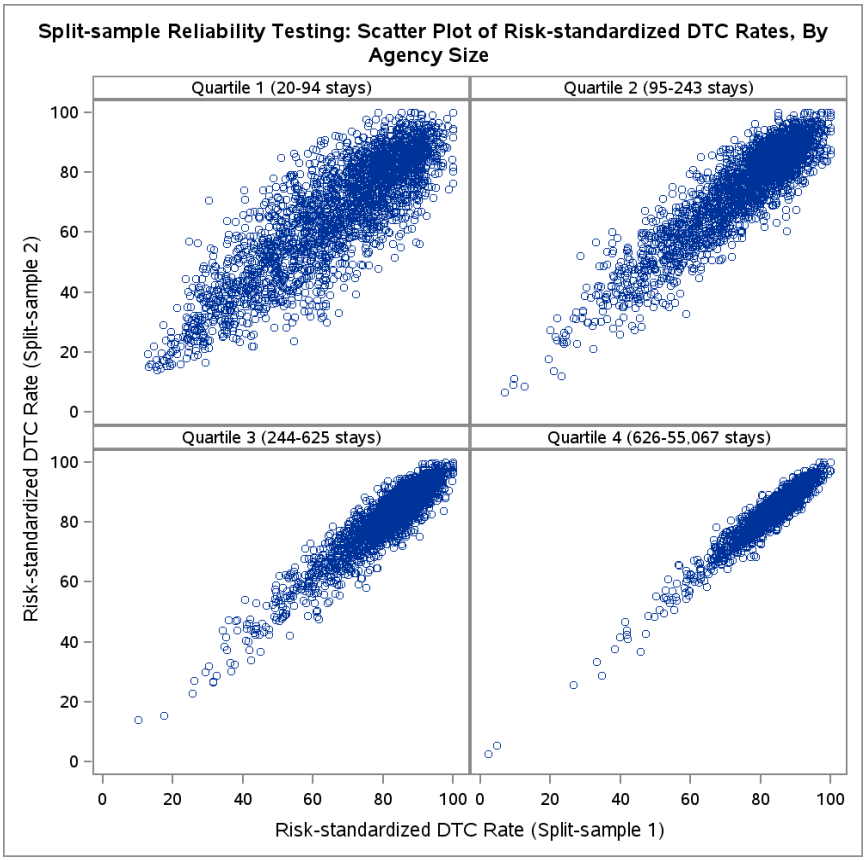
Acumen Analysis of Medicare Claims File for HH, CY 2015-2016. Note: The reliability testing was performed among HHAs with >= 20 stays. As a result, 1,543 HHAs were excluded from the analysis due to insufficient sample size.

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



Acumen Analysis of Medicare Claims File for HH, CY 2015-2016. Note: The reliability testing was performed among HHAs with >= 20 stays. As a result, 1,543 HHAs were excluded from the analysis due to insufficient sample size.

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



Acumen Analysis of Medicare Claims File for HH, CY 2015-2016. Note: The reliability testing was performed among HHAs with >= 20 stays. As a result, 1,543 HHAs were excluded from the analysis due to insufficient sample 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 strong in the overall sample, with the signal being 9.7 times as strong as the noise, on average. In the lowest quartile of HHAs the ratio is lower, with the signal bring 2.9 times as strong as the noise, which is still precise. This indicates that the measure score has good 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 HH sample was 0.91, indicating excellent reliability. When examined by facility sample size, the smallest quartile had good reliability and the second through fourth quartiles had excellent reliability.

2b1. VALIDITY TESTING

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**

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

☒ **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 (i.e., a higher score on a quality measure reflects higher quality).[10]

1. **Critical Data Elements Validity Testing:** The discharge to community measure outcome is defined based on three components: (i) discharge to a community setting from the HHA, (ii) no unplanned acute or LTCH hospitalizations in the day of discharge and the 31 days following HHA discharge, and (iii) no death in the day of discharge and the 31 days following HH 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 hospitalizations and death in the 31-day post-discharge window is small; thus, the discharge to community component is the key outcome determinant. Discharge to a community which is defined based on the "Patient Discharge Status Code" reported on the Medicare FFS HH claim[11]; Discharges to home or self-care with or discharged to home or self-care with a planned acute care hospital readmission, based on Patient Discharge Status Codes 01 or 81 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 separate 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 HH claim (critical data element) as compared to discharge to community based on the OASIS assessment (authoritative source)**

      This testing included 6,154,523 HH stays that were successfully matched to OASIS assessments. The OASIS discharge to community definition was based on OASIS-C1 item M0100 "This Assessment is Currently Being Completed for the Following Reason" with value 9 (Discharge from Agency) being classified as discharge to community. Using the OASIS discharges to community as the authoritative source, we examined the percentage of OASIS discharges to community that were also classified as discharged to community based on the HH claim.[12]

[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.

[12] M0100 Discharge from Agency also includes home health episodes ending with discharge to a home/community hospice. Discharge to hospice is an exclusion criterion for the measure denominator.

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

This testing was conducted on the total index HH sample of 6,279,286 patient stays. For patient stays classified as discharged to a community setting based on the HH claim, we looked for any acute, LTCH, skilled nursing facility (SNF), or IRF claims on day of or day after HH discharge. Absence of any claims would confirm the discharge to community reported on the HH 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 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.[13, 14, 15]

c. Known-groups validity testing

The known-groups validity method evaluates the degree to which a measurement demonstrates different scores for groups known to vary on the construct being measured.[16] We conducted known-groups validity testing by examining whether successful discharge to community rates varied for patient subgroups expected to have different likelihood of discharge to a community setting based on their clinical status at HH discharge.

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.

We also conducted patient-level validity tests, using a sub-sample of stays from CY 2015-2016 as a test of known group validity. We calculated the percentage of patients who were successfully discharged to the community as defined by the measure by several patient characteristics available in the OASIS-C1 assessment.

- **Known-groups validity testing for fallers versus non-fallers:** Using OASIS item M2310 (Reason for Emergent Care) response 2 -Injury caused by fall, we compared the successfully discharge to the community rate between patients who sought emergent care because of a fall and those who did not[17]. We hypothesized that patients/residents who sought emergent care because of a fall are less frequently successfully discharged to the community compared to those who did not.
- **Known-groups validity testing for patient cognitive function:** We also compared the successful discharge to community rate for the patient's level of alertness, orientation, comprehension, concentration, and immediate memory for simple commands at discharge from agency.
- **Known-groups validity testing for patients with varying functional abilities with self-care and mobility activities at discharge**

  We also examined the relationship between the activities of daily living and 11 mobility activities at discharge. The M1800-M1890 items on the OASIS-C1 assessment were used as a proxy for Section GG functional status items. These items are not completed for all patients discharged from an HHA (e.g.

---

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

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

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

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

[17] Patients who did not seek emergent care because of a fall or did not fall.

patients whose episode ends in a transfer to an inpatient facility or death at home). Therefore, this analysis focused on stays where the discharge functional codes of most able to least able were used. It is for these stays where we would expect a relationship between functional status and discharge destination. We hypothesized that patients who are independent on functional activities are more frequently successfully discharged to the community and conversely that patients who are dependent on those functional activities are less frequently successfully discharged to the community.

2. Performance Measure Score Empirical Validity Testing:

To assess convergent validity, we examined whether a facility's performance and percentile rank on the successful discharge to community measure was correlated with its performance and percentile rank on another quality measure in the same clinically-related group. We examined whether a facility's successful discharge to community rate and its percentile rank on five claims-based measures: Acute Care Hospitalization, Emergency Department Use without Hospitalization, Rehospitalization During the First 30 Days of Home Health, Emergency Department Use without Hospital Readmission During the First 30 Days of Home Health, and Potentially Preventable 30-Day Post-Discharge Readmission Measure. We calculated the Pearson Correlations between performance on the measures and Spearman Correlations between facilities percentile rank.

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

For many individuals, the goals of post-acute care include optimizing functional improvement and independence, returning to their previous level of independence and functioning if possible, and avoiding institutionalization. Discharge to a community-based setting is an important goal for the majority of patients in PAC settings, making it a priority for measure development from the patient and family perspective. For patients receiving home health services, the goal is typically for patients to remain in the home setting at the end of their home health episode. 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.

The importance and validity of this measure was a quality indicator was discussed during the Discharge to Community TEP meetings convened by RTI International and Abt Associates during August, September, and October 2015, as well as during a July 2011 IRF Technical Expert Panel convened by RTI International.[24] TEP members agreed that discharge to community is an important measure concept, noting that it is frequently the ultimate goal of rehabilitation, and gathering and reporting this outcome could be a powerful motivator for staff and patients to work toward functional goals to achieve this outcome.

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

1. **Critical Data Elements Validity:**

   a. **Correctness of discharge to community setting based on HH claim (critical data element) as compared to discharge to community based on the OASIS assessment (authoritative source)**

      Of the 5,003,067 HH patients who had a discharge to the community on the OASIS-C1 assessment, 96.5% were also classified as discharged to the community based upon the claims discharge status (**Table 5**).

Table 5. Agreement between of Claims Discharge Status Code and the Discharge Status on OASIS-C1 Assessment

Total N	Discharged to Community: Assessment	Discharged to Community: Claims	N agreement	% Agreement
6,154,523	5,003,067	5,271,256	4,827,412	96.5%

Acumen Analysis of Medicare Claims File and OASIS Assessments for HH, CY 2015-2016.

* Assessment discharge to community is defined by OASIS Assessment reason for assessment = 09 (Discharge from Agency).

** Claims Discharge to Community is defined by claims discharge code = 01 or 81 (discharge to home/self care [routine discharge]; discharged to home or self-care with a planned acute care hospital readmission).

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

      Only 1.6% of HH stays that were discharged to the community setting as noted on the HH claim had an acute, IRF, SNF, or LTCH claim on day of or day after HH discharge.

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

      In the merged datasets between HH claims and the OASIS-C1 assessments for the M2310 (Reason for Emergent Care) item, the successful DTC rate for CY 2015-2016 was 81.9% (**Table 6**). 1.3% of the sample sought emergent care because of a fall. In CY2015-2016, the successful DTC rate was significantly lower for patients who sought emergent care because of a fall (39.6%) compared to those who did not seek emergent care due to a fall or did not fall (82.5%).

[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

Table 6. Successful Discharge to the Community Rates: Differences between Patient Groups – Falls (Stay-level)

HH	Patients who Sought Emergent Care because of a Fall	All Other Patients[25]	All Patients
Successful Discharge to Community	29,732 (39.6%)	4,585,274 (82.5%)	4,615,006 (81.9%)
Total	75,117 (100.0%)	5,559,065 (100.0%)	5,634,182 (100.0%)

Acumen Analysis of Medicare Claims File and OASIS Assessments (M2310 Item) for HH, CY 2015-2016. Note: Analysis includes the Start of Care, Resumption of Care, and End of Care assessments. The analysis aggregates results from all assessments matching the stay from the first assessment to the last assessment.

d. Known-groups validity testing for patient cognitive function

The successful discharge to community rate was overall high for the alert and oriented (91.8%) and lower for totally dependent cognition function score at discharge from agency (85.7% - 89.3%) (**Table 7**).

Table 7. Successful DTC Rate by Cognitive Function Score at Discharge (Stay-level)

HH Cognitive Function Score	Count	Percent
Best: 0. Alert and Oriented	3,164,109	91.8%
1. Requires Prompting	994,451	89.3%
2. Requires Some Assistance	274,906	87.9%
3. Requires Considerable Assistance	88,847	86.8%
Worst: 4. Totally Dependent	20,187	85.7%

Acumen Analysis of Medicare Claims File and OASIS Assessments (OASIS-C1 item M1700 Cognitive Functioning Item at discharge from agency) for HH, CY 2015-2016.

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

Using OASIS-C1 ADL/IADL items M1800-M1890, we also examined the association between HH patients' discharge functional abilities and successful discharge to the community as defined in this measure. We hypothesized that patients who achieved higher discharge scores for each functional activity would be more likely to have a successful discharge to community. **Tables 8-1 and 8-2** show the percent of patients who had a successful discharge to community by discharge score, which ranges from most able to least able. A very high percentage of patients coded as "Most Able" (independent; highest functional ability) at discharge had a successful discharge to community (91.7% to 93.4%) whereas patients coded as "Least Able" (dependent; lowest functional ability) for each ability were less likely to be discharged to community successfully (76% to 87.2%). Within each item the pattern is generally monotonic: patients who were more independent were more likely to be discharged to community. Note that not all items have the same number of possible responses.

[25] Patients who did not seek emergent care because of a fall or did not fall.

Table 8-1. Successful DTC Rate by Activities of Daily Living Scores at Discharge (Stay-Level)

HH	Grooming (M1800)		Upper Body Dressing (M1810)		Lower Body Dressing (M1820)		Bathing (M1830)		Toilet Transferring (M1840)	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
Most Able (Fully Independent)	2,807,841	92.7%	2,723,422	92.8%	2,453,465	92.9%	1,049,455	93.4%	3,209,789	92.3%
	1,238,755	89.1%	1,243,526	89.2%	1,113,237	90.1%	1,395,803	92.6%	1,095,062	88.2%
	383,388	85.9%	460,756	86.3%	797,807	87.5%	1,106,766	90.0%	98,227	84.8%
	112,516	84.7%	114,796	84.7%	177,991	85.0%	675,249	87.5%	20,553	84.3%
							104,099	88.3%	118,869	84.6%
							109,635	85.5%		
Least Able (Dependent)							101,493	84.5%		

Acumen Analysis of Medicare Claims File and OASIS Assessments (M1800-M1890 Items) for HH, CY 2015-2016.

Table 8-2. Successful DTC Rate by Activities of Daily Living Scores at Discharge (Stay-Level)

HH	Toilet Hygiene (M1845)		Transferring (M1850)		Ambulation/Locomotion (M1860)		Feeding or Eating (M1870)		Meal Preparation (M1880)		Telephone Use (M1890)	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
Most Able (Fully Independent)	2,961,741	92.5%	2,155,574	92.9%	1,030,618	93.1%	3,336,954	92.0%	2,542,804	92.9%	3,456,213	91.7%
	1,051,260	89.1%	2,046,319	89.9%	1,240,514	92.6%	1,109,656	88.1%	1,250,429	89.3%	316,024	89.7%
	383,984	86.3%	237,806	85.3%	1,672,463	90.0%	74,178	85.7%	749,267	87.2%	280,577	88.6%
	145,515	85.1%	72,610	84.7%	389,998	86.7%	7,130	84.9%			175,355	87.9%
			9,684	83.1%	112,294	86.6%	12,751	83.2%			101,891	87.5%
			20,507	81.6%	81,897	85.3%	1,831	76.0%			117,962	87.0%
Least Able (Dependent)					14,716	80.5%						

Acumen Analysis of Medicare Claims File and OASIS Assessments (M1800-M1890 Items) for HH, CY 2015-2016.

2. Performance Measure Score Testing

We examined whether an HHA's performance on the DTC-PAC HH measure was correlated with its performance on another quality measure in the same clinically related group. Among HHA's that had sufficient denominator size for each measure, we calculated both the Pearson and Spearman correlations between the facility's measure performance the Discharge to Community measure and five claims based measures: Acute Care Hospitalization, Emergency Department Use without Hospitalization, Rehospitalization During the First 30 Days of Home Health, Emergency Department Use without Hospital Readmission During the First 30 Days of Home Health, and Potentially Preventable 30-Day Post-Discharge Readmissions.

As hypothesized, we found that the correlation between the HHA's measure performance scores and percentile ranks on the DTC measure and these five measures was negative, small, (r = -0.004 through -0.19), except for the PPR measure (r = -0.60) and statistically significant ($p < 0.001$) **Table 9**.

Table 9. Correlation between Quality Measures and DTC Expected Rate

Measure	Pearson Correlation	p	Spearman Correlation	p
Acute Care Hospitalization (NQF# 0171)	-0.18994	<.0001	-0.2102	<.0001
Emergency Department Use without Hospitalization (NQF# 0173)	-0.04831	<.0001	-0.0495	<.0001
Rehospitalization During the First 30 Days of Home Health (NQF# 2380)	-0.08026	<.0001	-0.08059	<.0001
Emergency Department Use without Hospital Readmission During the First 30 Days of Home Health (NQF#2505)	-0.00393	0.7848	0.00664	0.6443
Potentially Preventable 30-Day Post-Discharge Readmissions	-0.59861	<.0001	-0.59786	<.0001

Acumen Analysis of Medicare Claims File for HH, CY 2015-2016.

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; in addition to HH-specific literature, we reviewed literature related to LTCH, IRF, and SNF settings. Discharge to a community setting is an important goal for many HH 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.[26] 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.[27,28,29,30] In addition to being

[26] 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.

[27] 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.

[28] 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.

[29] Scheinhorn, D. J., Hassenpflug, M. S., Votto, J. J., Chao, D. C., Epstein, S. K., Doig, G. S., ... & Ventilation Outcomes Study Group. (2007). Post-ICU mechanical ventilation at 23 long-term care hospitals: a multicenter outcomes study. *Chest*, *131*(1), 85-93.

[30] 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.

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.[31,32]

Our TEP summary report is publicly posted.[33] The TEP members unanimously agreed that discharge to community is an important measure concept, noting that it is frequently the ultimate goal of rehabilitation, and gathering and reporting this outcome could be a powerful motivator for staff and patients to work toward functional goals to achieve this outcome. We also received several public comments during our call for public comment[34] and during the CY 2017 HH NPRM rulemaking process supporting the importance of the DTC-PAC HH QRP measure as a quality indicator in HH settings.[35]

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 Elements Validity Testing Results**
   a. **Correctness of discharge to community setting based on the HH claim:** Our findings indicated excellent accuracy of discharge to community based on the HH claim, as compared with both authoritative sources. Over 96% of patients identified as discharged to community based on the OASIS-C1 assessment were also identified as discharged to community based on the HH claim. Only 1.6% of patients identified as discharged to community based on the HH claim had an acute, IRF, SNF, or LTCH claim on the day of or day after HH discharge. The excellent agreement between OASIS-C1 assessment and HH claim discharge to community, along with the absence of institutional transfer claims in 98.4% of patients discharged to community provide strong evidence for critical data element validity of the HH claim discharge to community component based on the Patient Discharge Status code.
   b. **Known-groups validity**: At the patient stay-level the results presented above generally support the hypotheses that patients with higher functioning abilities, and patients not seeking emergent care for falls are more likely to be successfully discharged to the community. Taken together, these findings demonstrate that the critical data elements used to calculate the primary outcome for the Discharge to Community measure are correct and valid indicators of a patient's actual discharge destination.
2. **Performance Measure Score Validity Testing Results**: We examined whether an HHA score and rank on the Discharge to Community was correlated with its score and rank on other quality measures in the same clinically related group. We found that the correlation between an HHA's performance score

---

[31] 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.

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

[33] 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

[34] 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

[35] Medicare Program; CY 2017 Home Health Prospective Payment System Rate Update; Home Health Value-Based Purchasing Model; and Home Health Quality Reporting Requirements, 76702–76797 Available at: https://www.gpo.gov/fdsys/pkg/FR-2016-11-03/pdf/FR-2016-11-03.pdf

and percentile ranks on these were generally in the hypothesized direction, small, (r = -0.004 through -0.19), except for the PPR measure (r = -0.60) and statistically significant ($p < 0.001$). The relatively large magnitude of the correlation with the PPR measure was expected because the PPR measure captures post-discharge outcomes and requires a discharge to community/self-care status for inclusion in the measure denominator. With the Acute Care Hospitalization measure we would expect a smaller overlap in HH stays with the DTC measure, as it captures outcomes during the first 60 days of home health (i.e., the observation window can capture outcomes that may occur during the home health stay, a portion of the home health stay, or post-discharge) rather than solely capturing post-discharge outcomes. However, since both measures can capture post-discharge unplanned admissions, we would still expect to see a noticeable negative correlation between the two rates.

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

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 section **2b2.3**. We calculated the stay-level frequency of each exclusion and the HHA-level distribution of exclusions. Most exclusions were required to ensure availability of complete and valid data for measure specification (e.g., 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; thus, we did not examine the impact of these exclusions.

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, 22.7% of stays were excluded because of one or more exclusion criteria. The exclusion criteria with the greatest frequency is for patients who are not continuously enrolled in Medicare Fee-For-Service Parts A and B (not C) from 12 months prior to the start of the stay through the 31 days after discharge (12.1% excluded). The next most frequent reasons for exclusion were stays for patients enrolled in hospice during the post-discharge observation window (5.7%), stays for patients discharged to another home health agency (4.2%), and stays for patients that have a prior proximal stay with a principal diagnosis for medical, nonsurgical, treatment of cancer (2.0%).

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

Exclusion	N*	%
The patient is younger than 18 years old at the start of the stay.	87	0.0%
The patient is not continuously enrolled in Medicare Fee-For-Service Parts A and B (not C) from 12 months prior to the start of the stay through the 31 days after discharge. This exclusion only applies to months in which the patient is alive.	983,090	12.1%
The stay has problematic claims data (e.g. anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory).	18,340	0.2%

Exclusion	N*	%
The patient has a prior proximal stay (a short-term acute care or psychiatric stay within 30 days prior to home health admission) with a principal diagnosis for medical (nonsurgical) treatment of cancer.	163,926	2.0%
The patient is discharged to a psychiatric hospital.	0	0.0%
The patient is discharged against medical advice.	36,037	0.4%
The patient is discharged to a disaster alternative care site.	0	0.0%
The patient is discharged to hospice care.	151,578	1.9%
The patient is discharged to another home health agency.	344,689	4.2%
The patient was enrolled in hospice during the post-discharge observation window.	459,808	5.7%
The patient is discharged to a federal hospital.	1,090	0.0%
The patient is discharged to court/law enforcement.	0	0.0%
Total Number of Patient Stays Excluded	1,846,362	22.7%

* 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.

Acumen Analysis of Medicare Claims File for HH, CY 2015-2016.

Table 11 shows the distribution of exclusions at the agency-level. On average, 16.6% of an agency's stays were excluded because of patients who are not continuously enrolled in Medicare Fee-For-Service Parts A and B (not C) from 12 months prior to the start of the stay through the 31 days after discharge; 4.9% because the patients were enrolled in hospice during the post-discharge observation window; 4.9% because the patients are discharged to another home health agency, and 1.5% because the patients have a prior proximal stay with a principal diagnosis for medical, nonsurgical, treatment of cancer.

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

Exclusion	Mean	Min	25th	Median	75th	Max
The patient is younger than 18 years old at the start of the stay.	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
The patient is not continuously enrolled in Medicare Fee-For-Service Parts A and B (not C) from 12 months prior to the start of the stay through the 31 days after discharge. This exclusion only applies to months in which the patient is alive.	16.6%	0.0%	7.6%	12.4%	22.4%	100.0%
The stay has problematic claims data (e.g. anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory).	0.1%	0.0%	0.0%	0.0%	0.0%	86.4%
The patient has a prior proximal stay (a short-term acute care or psychiatric stay within 30 days prior to home health admission) with a principal diagnosis for medical (nonsurgical) treatment of cancer.	1.5%	0.0%	0.0%	1.0%	2.0%	100.0%
The patient is discharged to a psychiatric hospital.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
The patient is discharged against medical advice.	0.5%	0.0%	0.0%	0.0%	0.0%	100.0%
The patient is discharged to a disaster alternative care site.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
The patient is discharged to hospice care.	1.4%	0.0%	0.0%	0.3%	2.2%	50.0%
The patient is discharged to another home health agency.	4.9%	0.0%	2.3%	3.8%	5.8%	100.0%
The patient was enrolled in hospice during the post-discharge observation window.	4.9%	0.0%	2.1%	4.6%	6.8%	100.0%
The patient is discharged to a federal hospital.	0.0%	0.0%	0.0%	0.0%	0.0%	8.6%
The patient is discharged to court/law enforcement.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

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 12**, along with the rationale for each exclusion. About 12.1% of stays are excluded due to no continuous enrollment in Medicare Fee-For-Service Parts A and B (not C) from 12 months prior to the start of the stay through the 31 days after discharge. These patients are excluded because those without continuous Part A and B coverage, or those who are ever enrolled in a Part C Medicare Advantage plan during the pre- or post-PAC periods will not have complete Medicare FFS claims in the system thus incomplete information for determining outcome and risk adjustment. The second highest frequency of exclusion is for stays where the patient was enrolled in hospice during the post-discharge observation window (5.7%). 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.

Table 12. Exclusion Criteria and Rationale for the DTC-PAC HHA 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.
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 HH care plan may not have been fully implemented, and the discharge destination may not reflect the HHA's discharge recommendation. Additionally, patients discharged against medical advice may potentially be at higher risk of post-discharge admissions or death, depending on their medical condition, or due to potential non-adherence or non-compliance with care recommendations.
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 HH 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.
Patients discharged 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 provider, makes the final decision of discharge to hospice-home or hospice-facility.
Patients not continuously enrolled in Parts A and B FFS Medicare (or those enrolled in Part C Medicare Advantage) for the 12 months prior to the post-acute admission date, and at least 31 days after post-acute discharge date	Patients not continuously enrolled in Parts A and B FFS Medicare for the 12 months prior to the PAC admission date are excluded because risk adjustment for certain comorbidities requires information on acute inpatient bills for one year prior to post-acute admission. Patients not continuously enrolled in Part A FFS Medicare for at least 31 days after post-acute discharge are excluded because admissions and death must be observable in the 31-day post-discharge period. Patients without Part A and B coverage or those who are enrolled in Part C Medicare Advantage plans will not have complete inpatient claims in the system.
Patients whose prior short-term acute care stay was for non-surgical treatment of cancer	Patients whose 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 patients is consistent with the hospital-wide and post-acute readmission measures.
Post-acute stays that end in transfer to the same level of care	HHA stays that end in transfer to another HHA are excluded from the measure because the HHA episode has not ended. For a HHA episode that involves transfer to another HHA, only the final HHA provider is included in the measure.

Exclusion	Rationale
Post-acute 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 in the elements used for risk adjustment. No-pay post-acute stays involving exhaustion of Part A benefits are also excluded.
Medicare Part A benefits exhausted	Patients who have exhausted their Medicare Part A coverage during the PAC stay are excluded because the discharge destination decision may be related to exhaustion of benefits.

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 309 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. Patient characteristics related to discharge and a marker for the specific discharging HHA 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 HHA. 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 HHA effect, hierarchical modeling accounts for the known predictors of discharge to community, on average, such as patient characteristics, the observed facility rate, and the number of HHA stays eligible for inclusion in the measure. The estimated HHA effect is determined mostly by the HHA's own data if the number of patient discharges is relatively large (as the estimate would be relatively precise) but is adjusted toward the average if the number of patient discharges 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:

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

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

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 is a measure of the degree to which discharges to community are higher or lower than what would otherwise be expected. This standardized risk ratio is then multiplied by the mean discharge to community rate for all HHA stays for the measure, yielding the risk-standardized discharge to community rate for each HHA.

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

The risk model employs the following sets of covariates (See **attached excel document** for the full list of variables in the risk adjustment model):

(1) Demographics
   (a) Age and sex
   (b) Enrollment status
   (c) Activities of daily living scores

(2) Care received during the prior proximal hospitalization (if relevant)
   (a) Length of prior proximal hospitalization
   (b) Clinical classification software (CCS) diagnosis and procedure categories during prior proximal hospitalization

(3) Other care received within one year of the HH stay
   (a) Number of prior acute discharges
   (b) Number of outpatient emergency department visits
   (c) Number of skilled nursing facility visits
   (d) Number of long-term care hospital visits
   (e) Number of inpatient dialysis sessions
   (f) Hierarchical condition categories (HCC) comorbidities

Measure Calculation Algorithm

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

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

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

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

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

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

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

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

where the sum is over all stays in HHA_j, and ω_i is the random intercept.

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 HH-specific effect included in the predictions. This produces the expected number of discharges at the average HHA.

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 standardized risk ratio for each HHA, as the ratio of the predicted to expected number of discharges to community.

To calculate the HHA-wide standardized risk ratio, 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 HHA.

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

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

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

The DTC-PAC HHA measure is specific to HHA 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 HH 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 HH admission and during the prior acute hospitalization in the past 30 days, and certain comorbidities from acute care hospitalizations in the year preceding HH admission.

Variables such as race or ethnicity are often not included as risk 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 HHA; thus, we adjust for age-sex groups in our model.

Several risk adjusters used in our model were found useful in other 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. [36,37,38] 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. [39,40,41] Certain diagnoses including diabetes, stroke, Alzheimer's disease or dementia, psychiatric disorders, cancer, and end-stage renal disease, are associated with a decreased likelihood of

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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

[41] 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.

discharge to community.[42,43,44,45] Presence of comorbidities and severity of comorbidities has also been correlated with a decreased likelihood of discharge to community.[46,47,48]

Statistical approach for risk factor selection

Considering the volume of independent covariates available for this model, multiple steps were taken to eliminate variables that do not improve the model's predictive ability or that do not predict discharges to the community risk in a consistent manner. Variables flagged positively in less than 500 stays within the index HH stays were dropped as they are unlikely to meaningfully improve model performance. Furthermore, we eliminated all variables with zero discharges to the community in any individual year. This exclusion was necessary to run annual logistic regression models assessing each covariate's stability over time. To avoid overfitting these models, we used an additional logistic regression employing shrinkage to further subset the parameters to those that were most predictive of the DTC outcome. This regression was trained on a random eighty percent sample of the index HH stays. Lastly, we checked the remaining parameters for stability overtime by fitting a separate logistic model for each calendar year. If the point estimates for a particular variable were not consistently above or below the null across calendar years, then that variable was eliminated from the final model.

Our final model includes 309 risk adjustment variables: 23 age-sex interactions, four activity of daily living scores, 107 DGN flags, 56 HCCs, two original Medicare enrollment flags, one length of prior proximal hospitalization indicator, 10 number of prior acute discharge categories, 101 procedure code flags, and counts of prior ED visits, SNF visits, IRF visits, LTCH visits, and inpatient dialysis sessions.

---

[42] 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.

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

[44] 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

[45] 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.

[46] 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.

[47] 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

[48] 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)**

Data for rural location, and Medicaid dual status were readily available through the enrollment database (EDB) and Common Medicare Environment (CME) and analyzed during the measure testing process. While a scoping review[49] found general agreement that persons of lower socioeconomic status are not disadvantaged in terms of HH care services, there is a well-documented socioeconomic gradient seen with primary and acute care services. In addition, due to the availability of data and previous evidence on the importance of dual eligibility in similar claims-based measures, we tested the effects of patient-level dual eligibility status. For example, starting in fiscal year 2019, CMS's Hospital Readmission Reduction Program will assess penalties based on a hospital's performance relative to other hospitals treating a similar proportion of dual eligible patients.

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

See the **attached excel document** for the hierarchical logistic regression model results, including the specific risk adjusters used, and the corresponding coefficients, p-values, and odds ratios. The final model had a c-statistic of 0.730.

We conducted the following analysis for each social risk factor:

- Calculated the prevalence of patients with each social risk factor;
- Compared the unplanned admission rates following return to the community and successful discharge to community rates for patients with and without each social risk factor;
- Calculated the distribution of agency-level SDS factor prevalence;
- Calculated the difference in agency-level risk-standardized rates; and
- Assessed the difference in DTC measure scores estimated with and without dual eligibility adjustment for agencies 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 13. Comparison of Unplanned Admission Rates following Return to the Community and Successful Discharge to Community Rates, by Urban/Rural Status

Social Risk Factor	Total # of Stays	Stays with Successful Discharge to Community	Unadjusted DTC Rate	Stays with Unplanned Admission	Unplanned Admission Rate
Rural	1,394,248	1,064,728	76.4%	145,630	10.4%
Urban	4,884,941	3,861,610	79.1%	514,202	10.5%
Unknown	97	69	71.1%	14	14.4%
Total	6,279,286	4,926,407	78.5%	659,846	10.5%

Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

Table 14. Comparison of Unplanned Admission Rates following Return to the Community and Successful Discharge to Community Rates, by Dual Status

[49] Goodridge et al. Socioeconomic disparities in home health care service access and utilization: A scoping review 2012: *International J. Nursing Studies* 49(10); 1310-19

Social Risk Factor	Total # of Stays	Stays with Successful Discharge to Community	Unadjusted DTC Rate	Stays with Unplanned Admission	Unplanned Admission Rate
Full Dual	1,248,005	927,046	74.3%	154,482	12.4%
Dual without full Medicaid	383,448	283,365	73.9%	46,438	12.1%
Non-Dual	4,647,833	3,715,996	80.0%	458,926	9.9%
Total	6,279,286	4,926,407	78.5%	659,846	10.5%

Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

Table 15. Distribution of Agency-level SDS Factor Prevalence

SDS Factor	Mean	St.dev	Min	25th Pct*	Median	75th Pct	Max
Urban	76.20%	36.10%	0.00%	68.10%	96.70%	99.60%	100.00%
Rural	23.80%	36.10%	0.00%	0.40%	3.30%	31.90%	100.00%
Unknown Geography	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	6.70%
Full Dual	27.70%	22.10%	0.00%	11.50%	20.30%	38.50%	100.00%
Dual without full Medicaid	8.20%	8.70%	0.00%	2.20%	5.40%	11.30%	74.40%
Non-Dual	64.10%	23.60%	0.00%	48.80%	70.90%	82.20%	100.00%

* Pct = percentile

Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

Table 16. Effect of Dual Eligibility Status Indicators in the Hierarchical Logistic Regression Model

Dual Eligibility Indicator	Estimate	SE	*p*-value	OR	LCL	UCL
Full Dual	-0.1489	0.003	<.0001	0.862	0.857	0.867
Dual without full Medicaid	-0.07404	0.004	<.0001	0.929	0.921	0.937

Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

Table 17. Comparison of Agency-level Risk Standardized Rate (RSR)

	Mean	Std	Min	25th Pct*	Median	75th Pct	90th Pct	Max
Original Model	77.5%	16.8%	1.7%	71.6%	82.9%	88.8%	92.7%	100%
Model with Dual Eligibility Indicators	77.5%	16.8%	1.7%	71.7%	82.8%	88.7%	92.8%	100%
Agency-level Difference	0.4%	0.4%	0.0%	0.2%	0.3%	0.5%	0.9%	3.2%

* Pct = percentile

Acumen Analysis of Medicare Claims File for HH, Medicare Enrollment Database, Common Medicare Environment (CME) database, CY 2015-2016.

The unadjusted DTC rates and unplanned admission rates for urban/rural status were similar (**Table 13**). However, the rates differed for dual status. Specifically, patients who do not have dual eligibility status tend to have higher unadjusted DTC rates and lower unplanned admission rates (**Table 14**). Additionally, the prevalence of dual eligible patients differed amongst HHAs (**Table 15**). To further investigate the potential importance of these variables on agency-level risk adjustment, we re-ran the hierarchical logistical model including these variables and investigated the ensuing differences in the agency-level RSRs.

Table 16 shows the parameter estimates for the two dual eligibility status variables that were included. **Table 17** assesses the impact of including the dual eligibility status in the risk adjustment model. On average, there is a difference 0.4% in agency-level risk-standardized rates. The difference in agency-level risk-standardized rates for the 90th percentile are less than 1%. In addition, not shown in table, we found that controlling for these two dual eligible indicators increased the model C-statistic by only 0.003. Given the low impact of including dual eligibility status in the risk adjustment model, we do not recommend controlling for these social risk factors in the DTC 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*)

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

1. **Score Distribution**: We examined the distribution of the HHA-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 curve). The C-statistic is a measure of how accurately a statistical model can distinguish between a patient with and without an outcome. For binary outcomes the C-statistic is identical to the area under a ROC curve for the model. A C-statistic of 0.50 indicates random prediction, implying the model predicts no better than random 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 hierarchical logistic regression 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.

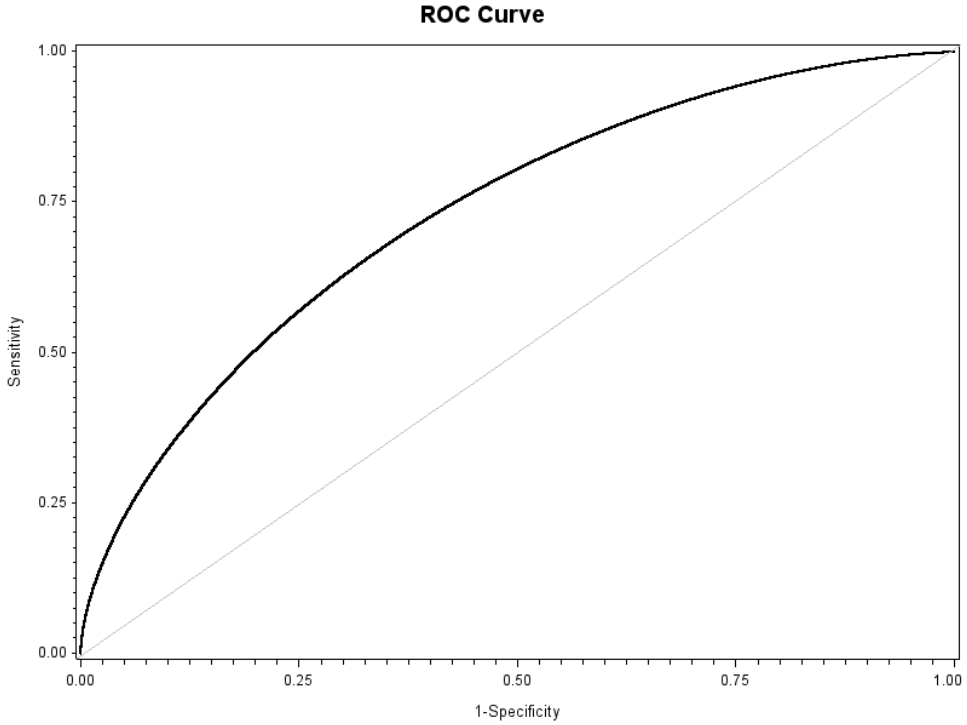
Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

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

The model C-statistic is 0.730 indicating that the model has good predictive ability. **Figure 2** below depicts the ROC curve for this model.

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

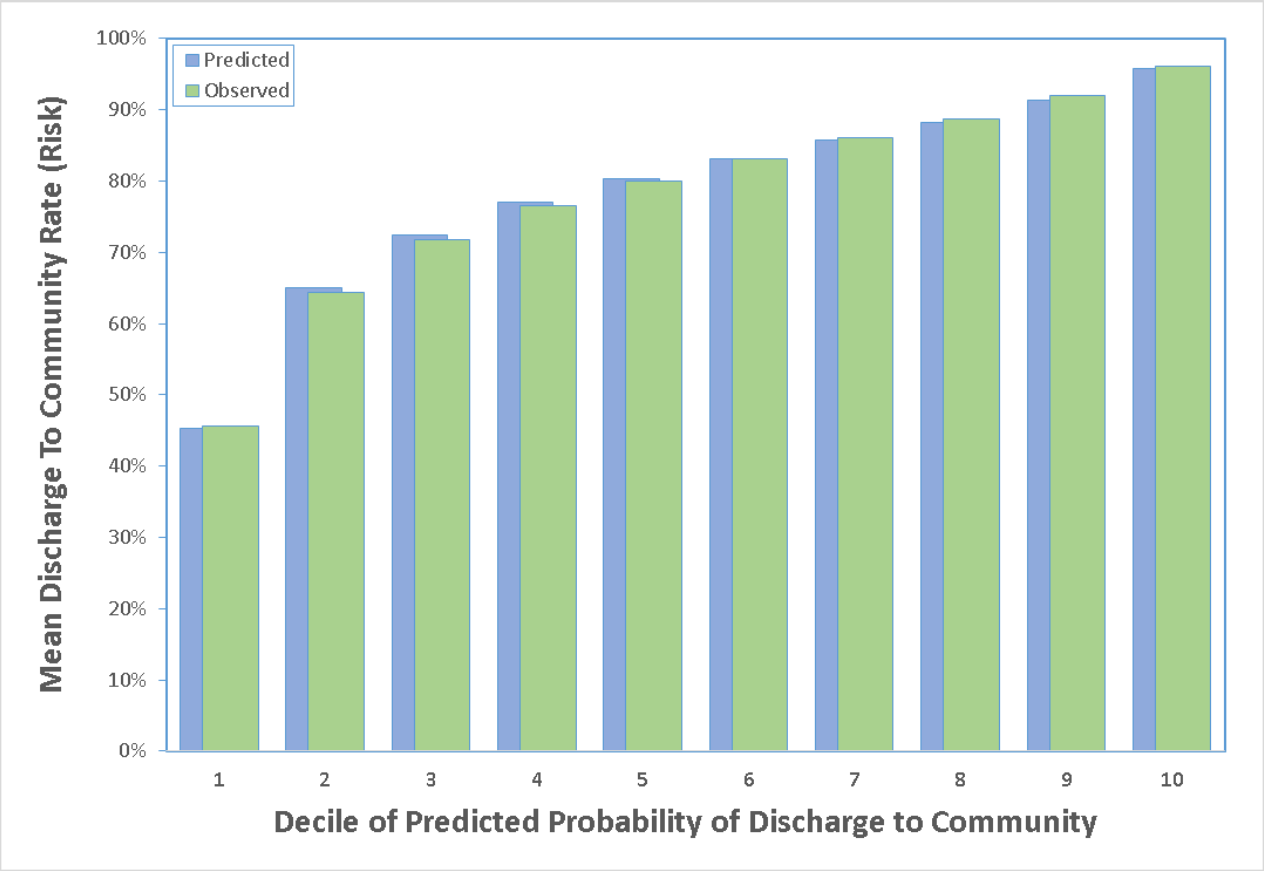


Acumen Analysis of Medicare Claims File for HH, CY 2015-2016.

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

The patient-level predicted Discharge to Community rates of the logistic model range from 1.7% to 100.0%, indicating that this model has a range of predictions and can predict both high and low scores. The results of our analysis comparing the observed and predicted DTC rates by deciles of DTC risk are displayed in **Figure 3** and **Table 18** In each risk decile, the observed and predicted DTC rates are close, with a difference of 0.7 percentage point or less. The ratio of observed to predicted rates is close to 1 across risk deciles, with the smallest ratio being 0.99 and the largest ratio being 1.01. This analysis is 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 one in this testing sample, as even minor deviations will come out as statistically significant. Our results demonstrate that this model demonstrates a good predictive ability to distinguish high-risk from low-risk subjects, thus demonstrating good calibration and discrimination.

Figure 3. HH Model Diagnostics: Comparison of Observed and Predicted Discharges to Community by Predicted Discharge to Community Deciles – CY 2015-2016



Acumen Analysis of Medicare Claims File for HH, CY 2015-2016.

Table 18. HH Model Diagnostics: Comparison of Observed and Predicted Discharges to Community by Predicted Discharge to Community Deciles – CY 2015-2016

Deciles of Predicted Probability of DTC	Number of HH Discharges	Observed Discharges to Community		Predicted Discharges to Community		Difference: Observed - Predicted		Ratio of Predicted DTC Rate/Observed DTC Rate
		Count	%	Count	%	Count	% points	
1	627,929	286,540	(45.6%)	284,654	(45.3%)	1,886	(0.3)	0.99
2	627,928	404,610	(64.4%)	408,378	(65.0%)	-3,768	(-0.6)	1.01
3	627,928	450,325	(71.7%)	454,689	(72.4%)	-4,364	(-0.7)	1.01
4	627,929	480,570	(76.5%)	483,555	(77.0%)	-2,985	(-0.5)	1.01
5	627,929	502,809	(80.1%)	504,861	(80.4%)	-2,052	(-0.3)	1.00
6	627,929	522,380	(83.2%)	522,490	(83.2%)	-110	(0.0)	1.00
7	627,927	540,361	(86.1%)	538,481	(85.8%)	1,880	(0.1)	1.00
8	627,928	557,634	(88.8%)	554,359	(88.3%)	3,275	(0.5)	0.99
9	627,930	577,747	(92.0%)	573,567	(91.3%)	4,180	(0.7)	0.99
10	627,929	603,431	(96.1%)	601,501	(95.8%)	1,930	(0.3)	1.00

Acumen Analysis of Medicare Claims File for HH, CY 2015-2016.

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

The results of our analysis comparing the observed and predicted DTC rates by deciles of DTC risk are displayed in **Figure 3** and **Table 18** above in section **2b3.7.**

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.** By taking into account beneficiary characteristics that are outside the provider's control, the model changes some providers' relative ranks of DTC rates. Risk adjustment compresses the distribution of DTC rates and decreases their variability. 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.

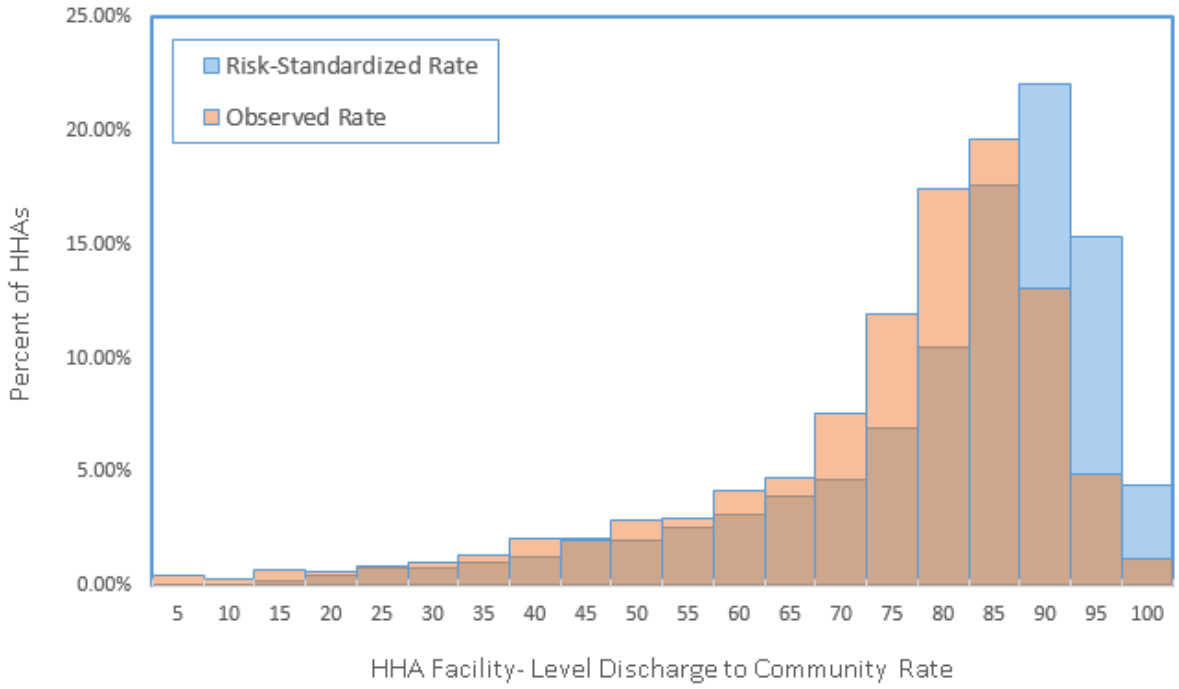
Table 19. Observed and Risk-standardized Discharge to Community Rates (%) among HHAs, CY 2015-2016

	K	Mean	SD	Min	25th Pct*	50th Pct	75th Pct	Max
Observed DTC rate	10,131	71.8%	17.6%	0.0%	65.5%	77.0%	83.5%	100.0%
Risk Standardized	10,131	77.5%	16.8%	1.7%	71.6%	82.9%	88.8%	100.0%

* Pct = percentile

Acumen Analysis of Medicare Claims File for HH, CY 2015-2016.

Figure 4. Observed and risk-adjusted discharge to community rates (%) among HHAs, CY 2015-2016



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*)

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 re-sampling and simulation techniques to assess the ability of the DTC rates to identify statistically significant differences in provider performance. This analysis was restricted to providers with 20 or more stays during CY 2015-2016 to align with the minimum sample size criterion for public reporting of the measure. More specifically, we used bootstrapping procedures to derive a confidence interval to determine if a HHA is performing better than, worse than, or no different than expected. Our bootstrapping approach aligns with that used in the all-caused hospital wide measure (HWR) measures.[50] We use bootstrapping methods across 500 bootstraps, sampling at the agency-level with replacement to generate the distribution of RSRs for each agency[51]. The 95% confidence interval (CI) is the range between the 2.5th and 97.5th percentile of each agency's distribution. We compared each agency's 95% CI to the national stay-level observed DTC rate (78.5%) to determine if the provider's performance was significantly different from the national rate. HHAs whose 95% CI was entirely below the national rate were considered to be significantly worse than the national rate; HHAs whose CI was entirely above the national rate were significantly better than the national rate; and HHAs whose CI overlapped the national rate were no different than 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*)

To determine if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified, we calculated the proportions of HHAs with scores that are significantly different from the national DTC rate (**Table 20**).

Table 20. Proportions of HHAs with Scores that are Significantly Different from the National DTC Rate

HHA Total	Number of facilities with 95% confidence interval lower than national mean (%)		Number of facilities with 95% confidence interval higher than national mean (%)		Total number of facilities with scores no different from mean (%)	
	K	%	K	%	K	%
10,131	2,434	24.0%	4,657	46.0%	3,040	30.0%

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 predicted Discharge to Community rates of the model range from 1.7% to 100%, indicating that this model has a range of predictions and can predict both high and low scores, given the distribution in the observed sample as well. These results regarding distribution and model discrimination indicated that this measure shows meaningful difference in our DTC risk standardized rate for individuals.

---

[50] Hospital-Wide All-Cause Readmission Measure (HWR) (CMS/Yale) (NQF #1789) Version7.0_Readmission_Hospital-Wide_Report_3-19-2018.pdf

https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/Hospital-Wide-All-Cause-Readmission-Updates.zip

[51] Consistent with the 2017 SAS pack for the 30-day hospital-wide readmission (HWR) measure, the calculation of the predicted rate within each bootstrap includes the addition of normally distributed noise proportional to the standard error of the provider-specific effect estimate.

Of the 10,131 HHAs, we found that most HHA's (70.0%) have scores for this measure that are significantly different from the national mean (**Table 20**). This supports the conclusions that there are meaningful differences in facility-level scores for this measure.

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, have only one set of specifications.

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, have only one set of specifications.

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, have only one set of specifications.

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 provider reimbursement, missing data are rare. Our measure excludes stays that are missing key measure specification data, under the exclusion criterion of claims with data that are problematic. 0.2% of stays were excluded from our measure due to problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory). Thus, missing data are rare and do not have an impact on the measure. Consequently, we do 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; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

This measure is calculated using Medicare FFS claims data; because submission and completion of claims is tied to provider reimbursement, missing data are rare. Our measure excludes stays that are missing key measure specification data, under the exclusion criterion of claims with data that are problematic. 0.2% of

stays were excluded from our measure due to problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory). Thus, missing data are rare and do not have an impact on the measure. Consequently, we do not perform any formal missing data analyses.

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; if no empirical analysis, 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).

Not applicable

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 FFS claims from the home health, inpatient, outpatient, and physician office settings 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, and 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 Home Health Quality Reporting Program https://www.medicare.gov/homehealthcompare/

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 will be publicly reported as part of the CMS's HH QRP.
Purpose:
The HH QRP creates HH quality reporting requirements, as mandated by Section 1895(b)(3)(B)(v)(II) of the Social Security Act (SSA) requiring the Secretary to establish quality reporting requirements for HHAs. More information about the HH QRP can be found at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/index.html. In addition to tracking quality of care, quality measure data are intended to help consumers make informed decisions when selecting healthcare providers.

Quality measure data from the HH QRP finalized for public reporting will be publicly reported at https://www.medicare.gov/homehealthcompare/ . HH quality measure data will also available for download for providers, researchers, and other public at https://data.medicare.gov/data/home-health-compare .

Geographic area and number and percentage of accountable entities and patients included:

The HH QRP includes all HHAs paid under the HH Prospective Payment System. 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 HHA measure results presented in this submission are based on 11,674 HHAs and 6,279,286 patient stays; these, data represent active HHAs with 20 or more eligible stays.

Level of measurement and setting

The DTC-PAC HHA measure is reported at the home health agency-level for HH 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?*)
NA - rollout for public reporting is underway. QM will be reported on Home Health Compare beginning January 2019.

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.*)

NA - rollout for public reporting is underway. QM will be reported Home Health Compare beginning January 2019.

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 HHA measure were provided to all active HH providers under the HH QRP in January, 2018. Active providers received provider preview reports in October, 2018, prior to public reporting of the DTC-PAC HHA measure in January 2019. Providers had a 30-day preview period to check these provider preview reports and submit suppression requests if there was evidence of errors in their data. 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 conducts open door forums during which stakeholders can ask general questions about the measure. CMS also conducted training sessions for home health providers to address the implementation of this measure. [15]. Along with the publicly-reported data, we also include consumer-friendly language to help consumers interpret measure data. Finally, DTC-PAC HHA measure specifications were publicly posted along with the CY17 HH QRP final rule at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/Proposed-Measure-Specifications-for-Measures-Proposed-in-CY-2017-HH-QRP-NPRM.pdf. The measure specifications are detailed and precise, allowing stakeholders to replicate measure calculations if they would like.

[15] Centers for Medicare & Medicaid Services Home Health Quality Reporting Program Medicare Learning Network - Implementing the IMPACT Act: Status Update & Activities. February 23, 2017.

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 HHAs that performed better than the national rate, number of HHAs that performed no different than the national rate, number of HHAs that performed worse than the national rate, and number of HHAs 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 HHA measure via a 30-day public comment period during November-December 2015, and during the CY 2017 HH QRP rulemaking process.

4a2.2.2. Summarize the feedback obtained from those being measured.

We received extensive support for implementation of the DTC-PAC HHA measure in the HH 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 CY17 HH QRP final rule with public comments and responses can be found at https://www.gpo.gov/fdsys/pkg/FR-2016-11-03/pdf/2016-26290.pdf.

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 FY17 HH QRP final rule with public comments and responses can be found at https://www.gpo.gov/fdsys/pkg/FR-2016-11-03/pdf/2016-26290.pdf.

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 Abt Associates Inc. reviewed and considered all public comments before the measure was finalized in the CY17 HH QRP 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 HHA measure was implemented on January 1, 2017 and will be publicly reported for the first time in January 2019 using CY 2015-2016 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.)

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_HHA_Appendix_2018-11-08-_FINAL.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): CMS - DCPAC

Co.2 Point of Contact: Tara, McMullen, McMullen, 410-786-8425-

Co.3 Measure Developer if different from Measure Steward: Abt Associates

Co.4 Point of Contact: Alrick, Edwards, Alrick_Edwards@abtassoc.com, 919-294-7735-

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 w 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

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

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: