Discharge Function Score for Home Health Agencies (HHAs)

Technical Report

June 2023

Submitted to:
Center for Clinical Standards and Quality
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244-1850

Contract No. 75FCMC18D0014, Task Order 75FCMC19F0001

Submitted by:
Abt Associates
6130 Executive Boulevard
Rockville, MD 20852

CONTENTS

TABLE OF EXHIBITS

1. Introduction

The Patient Protection and Affordable Care Act (ACA) of 2010[1] and Improving Medicare Post-Acute Care Transformation Act (IMPACT) of 2014[2] require the Secretary to establish public reporting requirements for quality measures for home health agencies (HHAs) using standardized patient assessment data elements. As part of this mandate, the Centers for Medicare & Medicaid Services (CMS) has contracted with Abt Associates to develop a cross-setting functional outcome measure to be used in the HH Quality Reporting Program (QRP) under the *Home Health and Hospice Quality Reporting Program Quality Measures and Assessment Instruments Development, Modification and Maintenance, & Quality Reporting Program Oversight Support* contract (75FCMC18D0014/Task Order 75FCMC19F0001).

Measuring functional status of HH patients can provide valuable information about an HHA's quality of care. A patient's functional status is associated with institutionalization,[3] higher risk of falls and falls-related hip fracture and death,[4,5] greater risk of undernutrition,[6] higher emergency department admissions,[7] higher risk of readmissions following home care,[8] and higher prevalence of hypertension and diabetes.[9] Predictors of poorer recovery in function

[1] Section 3004(b) of the Patient Protection and Affordable Care Act of 2010, Pub.L. 111-148

[2] Amendment Section 1899B to the Social Security Act, Pub.L. 113-185

[3] Hajek, A., Brettschneider, C., Lange, C., Posselt, T., Wiese, B., Steinmann, S., Weyerer, S., Werle, J., Pentzek, M., Fuchs, A., Stein, J., Luck, T., Bickel, H., Mösch, E., Wagner, M., Jessen, F., Maier, W., Scherer, M., Riedel-Heller, S.G., König, H.H., & AgeCoDe Study Group. (2015). Longitudinal Predictors of Institutionalization in Old Age. PLoS One, 10(12):e0144203.

[4] Akahane, M., Maeyashiki, A., Yoshihara, S., Tanaka, Y., & Imamura, T. (2016). Relationship between difficulties in daily activities and falling: loco-check as a self-assessment of fall risk. Interactive Journal of Medical Research, 5(2), e20.

[5] Zaslavsky, O., Zelber-Sagi, S., Gray, S. L., LaCroix, A. Z., Brunner, R. L., Wallace, R. B., … Woods, N. F. (2016). Comparison of Frailty Phenotypes for Prediction of Mortality, Incident Falls, and Hip Fracture in Older Women. Journal of the American Geriatrics Society, 64(9), 1858-–1862.

[6] van der Pols-Vijlbrief, R., Wijnhoven, H. A. H., Bosmans, J. E., Twisk, J. W. R., & Visser, M. (2016). Targeting the underlying causes of undernutrition. Cost-effectiveness of a multifactorial personalized intervention in community-dwelling older adults: A randomized controlled trial. Clinical Nutrition (Edinburgh, Scotland).

[7] Hominick, K., McLeod, V., & Rockwood, K. (2016). Characteristics of older adults admitted to hospital versus those discharged home, in emergency department patients referred to internal medicine. Canadian Geriatrics Journal : CGJ, 19(1), 9–14.

[8] Middleton, A. Downer, B., Haas, A., Knox, S., & Ottenbacher, K.J. (2019) Functional status ss associated with 30-day potentially preventable readmissions following home health Care. Medical Care, 57(2):145-151.

[9] Halaweh, H., Willen, C., Grimby-Ekman, A., & Svantesson, U. (2015). Physical activity and health-related quality of life among community dwelling elderly. J Clin Med Res, 7(11), 845–52.

include greater age, complications after hospital discharge, and residence in a nursing home.[10] Understanding factors associated with poorer functional recovery facilitates the ability to estimate expected functional outcome recovery for patients, based on their personal characteristics.

Home health care can positively impact functional outcomes. In stroke patients, home-based rehabilitation programs administered by home health clinicians significantly improved function.[11] Home health services, delivered by a registered nurse positivity impacted patient Quality of Life (QOL) and clinical outcomes, including significant improvement in dressing lower body and bathing activities of daily living, meal preparation, shopping, and housekeeping instrumental activities of daily living.[12] In addition, a retrospective study, using data abstracted from the Minimum Data Set and OASIS, reported that nursing home admissions were delayed in the study population receiving home health services by an average of eight months[13] and for a similar population, community dwelling adults receiving community-based services supporting aging in place, enhanced health and functional outcomes, improved cognition and lower rates of depression, function assistance, and incontinence were noted.[14]

The Discharge Function Score measure determines how successful each HHA is at achieving an expected level of functional ability for its patients at discharge. The measure was placed on the 2022 Measures Under Consideration (MUC) List as the "Cross-Setting Discharge Function Score," and the measure name was subsequently modified to the "Discharge Function Score." An expectation for discharge function score is built for each HHA episode by accounting for patient characteristics that impact their functional status. The final Discharge Function Score for a given HHA is the proportion of that HHA's episodes where a patient's observed discharge score meets or exceeds their expected discharge score. HHAs with low scores are not producing the functional gains that they could be for a larger share of their patients. The measure provides actionable feedback to HHAs that has the potential to hold providers accountable and encourage

---

[10] Córcoles-Jiménez, M. P., Villada-Munera, A., Del Egido-Fernandez, M. A., Candel-Parra, E., Moreno-Moreno, M., Jimenez-Sanchez, M. D., & Pina-Martinez, A. (2015). Recovery of activities of daily living among older people one year after hip fracture. Clinical Nursing Research, 24(6), 604–623.

[11] Asiri, F. Y., Marchetti, G. F., Ellis, J. L., Otis, L., Sparto, P. J., Watzlaf, V., & Whitney, S. L. (2014). Predictors of functional and gait outcomes for persons poststroke undergoing home-based rehabilitation. Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association, 23(7), 1856–1864.

[12] Han, S. J., Kim, H. K., Storfjell, J., & Kim, M. J. (2013). Clinical outcomes and quality of life of home health care patients. *Asian Nursing Research*, *7*(2), 53-60.

[13] Young, Y., Kalamaras, J., Kelly, L., Hornick, D., & Yucel, R. (2015). Is Aging in Place Delaying Nursing Home Admission? Journal of the American Medical Directors Association, 16(10), 900.e1–6.

[14] Marek, K.D., Popejoy, I., Petroski, G. et al. (2005). Clinical outcomes of aging in place. Nurs Res; 54:202–211.

them to improve the quality of care they deliver. This measure also promotes patient wellness, encourages the provision of adequate therapy to help prevent adverse outcomes (e.g., re-hospitalization), and increases the transparency of quality of care in the HH setting. The Discharge Function Score measure adds value to the HH QRP function measure portfolio by using specifications that allow for better comparisons across post-acute care (PAC) settings, considering both self-care and mobility activities in the function score, and refining the approach to addressing missing item scores.

Input from a variety of stakeholders has been taken into consideration throughout the measure development process. Feedback was sought and considered from patients and caregivers on the salience of the measure concept and from Technical Expert Panels (TEPs) on the appropriate specifications for the cross-setting measure.

This report presents the Discharge Function Score measure specifications. Section 2 provides an overview of the measure and is a high-level summary of the key features of the measure that are described in detail in the remaining sections of the document. Section 3 describes the methodology used to construct the Discharge Function Score measure including its data sources, study population, measure outcome, and steps for calculating the final measure score. Section 4 discusses Discharge Function Score measure testing, including the measure's reportability, variability, reliability, and validity testing results. Lastly, the Appendix includes risk adjustment model results and supporting information for the statistical imputation models used to estimate missing item scores.

2. Overview

This section provides an overview of basic descriptive information on the Discharge Function Score measure, summarizing the key points contained in the rest of the document. A more detailed explanation of the measure specifications is available in Section 3.

2.1 Measure Name

Discharge Function Score

2.2 Measure Type

Outcome Measure

2.3 Care Setting

HH

2.4 Data Source

Outcome and Assessment Information Set (OASIS)

2.5 Brief Description of Measure

The Discharge Function Score calculates the percent of HH patients who achieve a risk-adjusted expected function score at discharge. Functional status is measured through Section GG of OASIS assessments, which evaluates a patient's capacity to perform daily activities related to self-care (GG0130) and mobility (GG0170). Coefficients from a risk adjustment model controlling for admission function score, age, and patient clinical characteristics are used to determine an expected discharge function score for each HH episode. The provider score is calculated as the following proportion:

$$\frac{Number\ of\ HHA's\ episodes\ where\ observed\ \text{discharge score} \geq\ expected\ discharge\ score}{Total\ number\ of\ HHA's\ episodes} * 100$$

3. Measure Specifications

3.1 Measure Time Period

This measure is calculated using 12 months (four quarters) of data. All HH episodes with a discharge date that falls within this target period, except those that meet the exclusion criteria (refer to Section 3.3.2 for details), are included in the measure.

3.2 Data Source

This measure uses data from the OASIS. The OASIS data are collected on all Medicare patients who receive services from an HHA. This measure is calculated entirely using administrative data. There will be no additional data collection or submission burden for HHAs.

3.3 Denominator

The denominator is the total number of HH episodes with an OASIS record in the measure target period, which do not meet the exclusion criteria.

3.3.1 HH Episode Construction

We use HH episodes to construct this measure. The target date for an HH episode is the end of care (EOC) date. The target period for the measure is 12 months (four quarters). Documentation on how HH episodes are constructed is available in the Home Health Quality Reporting Program Measure Calculations and Reporting User's Manual: Version 2.0[15].

3.3.2 Eligible Episodes

The eligible episodes for this measure are all HH episodes that do not meet the exclusion criteria during the target period. The HH episode is excluded if any of the following are true:

- Patients with an incomplete stay. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients who die; and patients with an HH episode that is less than 3 days.

  Rationale: When a patient has an incomplete stay, for example, the patients leave urgently due to a medical emergency, it can be challenging to gather accurate discharge functional status data.

- Patient is in a coma, persistent vegetative state, has complete tetraplegia, locked-in state, severe anoxic brain damage, cerebral edema, or compression of the brain.

[15] https://www.cms.gov/files/document/hh-qrp-qm-users-manual-v20pdf.pdf

Rationale: These patients are excluded because they may have limited or less predictable mobility improvement with the selected items.

- Patient is younger than 18 years: Age in years is calculated based on the truncated difference between admission date and birth date, i.e., the difference is not rounded to nearest whole number.

  Rationale: Patient under 18 years old are not included in the target population for this measure because pediatric HH patients may have different patterns of care than adult patients.

- Patient is discharged to hospice (home or institutional facility).

  Rationale: Patient goals may change during the HH episode, and functional improvement may no longer be a goal for a patient discharged to hospice.

3.4 Numerator

The numerator is the number of episodes in an HHA with an observed discharge function score (Section 3.4.1) for Section GG function items that is equal to or higher than the calculated expected discharge function score (Section 3.4.2).

3.4.1 Observed Discharge Function Score

The observed discharge function score is the sum of individual function items at discharge. Section GG of each PAC assessment instrument includes standardized patient assessment data elements that measure functional status. The Discharge Function Score measure focuses on GG items that are currently available across these PAC settings (Exhibit 1).

Exhibit 1. Cross-Setting Function Item Set

Item	Item Description
GG0130A	Eating
GG0130B	Oral Hygiene
GG0130C	Toileting Hygiene
GG0170A	Roll Left and Right
GG0170C	Lying to Sitting on Side
GG0170D	Sit to Stand
GG0170E	Chair/Bed-to-Chair Transfer
GG0170F	Toilet Transfer
GG0170I	Walk 10 Feet
GG0170J	Walk 50 Feet with 2 Turns
GG0170R	Wheel 50 Feet with 2 Turns

Valid responses for GG items are reported in Exhibit 2.

Exhibit 2. GG Items Response

Category	GG Items Response	Response Description
Patient Functional Status Assessed	6	Independent
	5	Setup or clean-up assistance
	4	Supervision or touching assistance
	3	Partial/moderate assistance
	2	Substantial/maximal assistance
	1	Dependent
Activity Not Attempted (ANA) codes	7	Patient refused
	9	Not applicable
	10	Not attempted due to environmental limitations
	88	Not attempted due to medical condition or safety concerns
Other NA codes	^	Skip pattern
	-	Not assessed/no information

The following steps are used to determine the observed discharge function score for each episode:

Step 1: If the code for an item is between 1 and 6, then use code as the score for that item.

Step 2: If code for an item is 7, 9, 10, 88, dashed (-), skipped (^), or missing, then use statistical imputation to estimate the item score for that item (see Section 3.5).

Step 3: Sum scores across all items to calculate the total observed discharge function score. Different locomotion items are used if the patient uses a wheelchair than for the remaining patients.

Use 2 * Wheel 50 Feet with 2 Turns (GG0170R) score to calculate the total observed discharge function score for episodes where (i) Walk 10 Feet (GG0170I) has an activity not attempted (ANA) code at both admission and discharge and (ii) either Wheel 50 Feet with 2 Turns (GG0170R) or Wheel 150 Feet (GG0170S) has a code between 1 and 6 at either admission or discharge. The remaining episodes use Walk 10 Feet (GG0170I) + Walk 50 Feet with 2 Turns (GG0170J) to calculate the total observed discharge function score.

In either case, 10 items are used to calculate a patient's total observed discharge score and score values range from 10 – 60.

3.4.2 Expected Discharge Function Score

The expected discharge function score is determined by applying the regression equation determined from risk adjustment to each HH episode. Risk adjustment controls for resident characteristics such as admission function score, age, and clinical conditions. Refer to Section 3.6 for details on risk adjustment.

3.5 Statistical Imputation

When an item score is missing because an ANA code, a dash (-), or a skip (^) has been recorded (henceforth referred to as NA) rather than a value of 1 to 6, item scores are estimated through statistical imputation. This approach refines the imputation method used for in-use IRF QRP Functional Outcome Measures: Change in Self-Care Score for Medical Rehabilitation Patients (CBE ID #2633), Discharge Self-Care Score for Medical Rehabilitation Patients (CBE ID #2635), Change in Mobility Score for Medical Rehabilitation Patients (CBE ID #2634), and Discharge Mobility Score for Medical Rehabilitation Patients (CBE ID #2636), which were measures being adapted for HH QRP.[16] The method used in these measures recodes all NAs to 1, which implicitly assumes all NA codes signify residents who are completely dependent on a functional activity. On average, patients who are coded as NA on a GG item at admission tend to score higher at discharge (if assessed) than patients who are coded as dependent on admission. Treating both types of patients the same in risk adjustment can lead to less accurate expected discharge values for each of these types of patients. Statistical imputation allows NAs to take any value from 1 to 6, based on a patient's clinical characteristics and codes assigned on other GG items.

A separate statistical imputation model is constructed for each GG item used in the Discharge Function Score (Section 3.4.1) at admission and discharge. Imputation models include the predictors used in risk adjustment (Section 3.6.2) and covariates for scores on other GG items (Step 3 below). Notably, imputation models use all GG items available in HH to estimate missing scores for the subset of GG items used for the Discharge Function Score numerator (detailed imputation model results are available upon request). The following steps are used to generate imputed item scores for episodes with NA codes. Note that these steps first describe imputing a single item at admission and then describe the relevant modifications for discharge and for the other items.

---

[16] Consensus-Based Entity (CBE), as noted in Section 1890 of the Social Security Act. Section 1890 of the Social Security Act requires the Secretary of HHS to contract with a CBE regarding performance measurement. The National Quality Forum (NQF) was the CBE from 2010 – 2023. Battelle Memorial Institute has been contracted as the CBE from March 2023—March 2028. In this rule and henceforth, references to NQF will be replaced with CBE.

Step 1: Start with Eating (GG0130A). Identify eligible episodes where the item score is not missing (i.e., had a score 1-6) at admission. These scores are used as the outcome (i.e., left-hand-side variable) of the admission imputation model for GG0130A.

Step 2: For each episode, determine whether to use walking or wheeling items in the imputation model.

a) If Walk 10 Feet (GG0170I) has an ANA code at both admission and discharge and either Wheel 50 Feet with 2 Turns (GG0170R) or Wheel 150 Feet (GG0170S) has a code between 1 and 6, then use wheeling items.

b) Otherwise, use walking items.

Step 3: Create variables for the imputation model reflecting how each item (g_2 through g_{10}) except Eating (GG0130A) was scored at admission. GG item scores are described as independent variables (i.e., on the right-hand side) by three variables, collectively referred to as g'. The first reflects a score of 1-6 when available (g), the second is an indicator variable taking a value of 1 if the item had an ANA code, dash, or missing value (g^*), and the third is an indicator variable taking a value of 1 if the item was skipped (g^{**}).

$$Function\ items: G \in \{g_2, \dots, g_{10}\} \quad (1)$$

$$g' = [g, g^*, g^{**}] \quad (2)$$

$$g = \begin{cases} g, & g = \{1,2,3,4,5,6\} \\ 0, & otherwise \end{cases}$$

$$g^* = \begin{cases} 1, & g = \{7,9,10,88,-\} \\ 0, & otherwise \end{cases}$$

$$g^{**} = \begin{cases} 1, & g = \{\hat{}\} \\ 0, & otherwise \end{cases}$$

$$Function\ items\ with\ NA\ indicators: G' \in \{g'_2, \dots, g'_{10}\} \quad (3)$$

Step 4: Estimate an ordered probit model using the sample identified in Step 1.

Two types of predictors (i.e., right-hand-side variables) are used in the imputation method: clinical covariates (C) and function items with NA indicators (G') constructed in Step 3.

$$Clinical\ items := C \in \{c_1, \dots, c_k\} \quad (4)$$

$$Function\ items\ with\ NA\ indicators: G' \in \{g'_2, \dots, g'_{10}\} \quad (5)$$

The model we estimate for g_1, GG0130A, is

$$z_i = C_i\beta + G_i'\phi + \varepsilon_i \tag{6}$$

$$g_i = \begin{cases} 1, & z_i \le \alpha_1 \\ 2, & \alpha_1 < z_i \le \alpha_2 \\ 3, & \alpha_2 < z_i \le \alpha_3 \\ 4, & \alpha_3 < z_i \le \alpha_4 \\ 5, & \alpha_4 < z_i \le \alpha_5 \\ 6, & z_i > \alpha_5 \end{cases} \tag{7}$$

The latent variable, z_i, is interpreted as patient i′s underlying degree of independence on assessment item GG0130A and is a continuous variable. The error term, ε_i, is assumed to be independent and identically distributed N(0,1). The model assumes that the assessment item, g_i, because it only can take on six levels, discretizes the underlying continuous independence. It does this using thresholds: patients whose underlying independence is lower than the lowest threshold, α_1, are coded as most dependent and given a score of 1; patients whose level of dependence is a bit higher, higher than the lowest threshold α_1 but lower than the second lowest threshold α_2, achieve a score of 2 on this item. This proceeds until we are considering patients whose independence is higher than the highest threshold, α_5, who receive a score of 6.

We compute the imputed value of g_i as

$$\begin{aligned} \hat{g}_i = {} & \Pr(z_i \le \alpha_1) + 2 * \Pr(\alpha_1 < z_i \le \alpha_2) + 3 * \Pr(\alpha_2 < z_i \le \alpha_3) + \\ & 4 * \Pr(\alpha_3 < z_i \le \alpha_4) + 5 * \Pr(\alpha_4 < z_i \le \alpha_5) + 6 * \Pr(z_i > \alpha_5) \end{aligned} \tag{8}$$

Step 5: Repeat Steps 1 - 4 for Eating (GG0130A) at discharge, replacing the word "admission" with the word "discharge" in Steps 1 - 4.

Step 6: Repeat Steps 1 - 5 for each GG item included in the observed discharge function score (Section 3.4.1), as above replacing the Eating (GG0130A) item with each successive GG item in Steps 1-5. For Wheel 50 Feet with 2 Turns (GG0170R), use only the sample of episodes that satisfies the conditions in Step 2a. For Walk 10 Feet (GG0170I) and Walk 50 Feet with 2 Turns (GG0170J), use only the sample of episodes that satisfies the conditions in Step 2b.

3.6 Risk Adjustment

The purpose of risk adjustment is to account for differences across HH patients that affect their functional status. Risk adjustment creates an individualized expectation for discharge function score for each episode that controls for admission functional status, age, and clinical characteristics. This ensures that each episode is measured against an expectation that is calibrated to the patient's individual circumstances when determining the numerator for each HHA. See Exhibit A-1 for risk adjustment model results.

3.6.1 Statistical Risk Model

The statistical risk model is an ordinary least squares linear regression model, which estimates the relationship between discharge function score and a set of risk adjustors. Observed discharge function score is determined for each HHA episode, incorporating imputed item scores when NA codes are encountered. The risk adjustment model is run on all HHA episodes to determine the model intercept (β_0) and risk adjustor coefficients ($\beta_1, \ldots, \beta_n$). Expected discharge function scores are calculated by applying the regression equation to each HHA episode.

$$Expected\ Discharge\ Function\ Score = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \qquad (9)$$

where $x_1 - x_n$ are the risk adjustors.

3.6.2 Variables

This section contains a listing of covariate groups used to calculate the risk-adjusted discharge function scores. Information on the covariates were obtained from the OASIS data.

- Age Category

  Age was calculated as the difference between the start of care/resumption of care (SOC/ROC) date of the HH episode and the patient's date of birth.

- Admission Function Score

  Admission function score is the sum of admission scores for function items included in the discharge score (Section 3.4.1). It can range from 10-60, with a higher score indicating greater independence. NAs in the admission item scores are treated the same way as NAs in the discharge item scores, with NAs replaced with imputed scores (Steps 1-2 in Section 3.4.1). The walking items and wheeling item are used in the same manner as in the discharge score (Step 3 in Section 3.4.1). Admission score squared is also included as a risk adjustor.

- Prior surgery

  This covariate captures whether the patient had prior surgery.

- Prior Function/Device Use

  These covariates capture patient's functional status prior to the episode.

- Pressure Ulcers

  These covariates capture the presence of pressure ulcer(s) at different stages.

- Cognitive Function

  These covariates capture the patient's cognitive function by assessing whether the resident's mental status at admission is impaired, and if impaired, at what level.

- Incontinence

  These covariates indicate the patient's level of bladder and bowel incontinence.

- Availability of Assistance and Living Arrangements

  These covariates indicate the patient's residential circumstance and availability of assistance.

- Admission Source

  These covariates indicate whether the patient was admitted from the community at SOC or from a facility at SOC/ROC.

- Body Mass Index

  These covariates indicate whether the patient has a low BMI ($12 \leq BMI \leq 19$) or high BMI (>50).

- Risk for hospitalization

  These covariates indicate a history of falls, multiple hospitalizations, multiple ER visits, decline in status, non-compliance, or polypharmacy prior to SOC/ROC.

- Confusion

  These covariates indicate whether the patient has moderately frequent or severely frequent confusion in the 14 days prior to SOC/ROC.

- Vision

  These covariates indicate whether the patient has partial or severely impaired vision.

- Medication Management Needs

  These covariates indicate whether the patient needs medication management assistance for oral or injectable medication.

- Supervision and Safety Sources of Assistance

  These covariates indicate whether the patient needs and has non-agency caregivers with proper training.

- HCC Comorbidities

  Comorbidities are obtained from Items M1021 and M1023 in OASIS. Comorbidities are grouped using CMS Hierarchical Condition Categories (HCC) software version 24.

3.7 Measure Calculation

The Cross-Setting Discharge Function Score is the proportion of HH episodes where residents achieve an expected discharge function score at discharge. A higher score indicates better performance in functional outcomes. For each HH episode, observed discharge function score (Section 3.4.1) and expected discharge function score (Section 3.4.2) are determined. For each HHA, the Cross-Setting Discharge Function Score is the proportion of episodes where the observed discharge function score is larger than or equal to the risk-adjusted expected function score.

3.7.1 Steps Used in Calculation

Step1: Calculate the observed discharge function score as described in Section 3.4.1, incorporating imputed item scores (Section 3.5).

Step 2: Identify excluded HH episodes using the criteria mentioned in Section 3.3.2.

Step 3: Calculate the expected discharge function score. For each HH episode: use the intercept and regression coefficients to calculate the expected discharge function score using the formula mentioned in Section 3.6. Note that any expected discharge function score greater than the maximum (i.e., 60) would be recoded to the maximum score.

Step 4: Calculate the difference in observed and expected discharge function scores. For each HH episode which does not meet the exclusion criteria, compare each resident's observed discharge function score (Step 1) and expected discharge function score (Step 3) and classify the difference as one of the following:

- Observed discharge score is equal to or higher than the expected discharge score.
- Observed discharge score is lower than the expected discharge score.

Step 5: Determine the denominator count. Determine the total number of HH episodes with an OASIS target date in the measure target period, which do not meet the exclusion criteria.

Step 6: Determine the numerator count. The numerator for this quality measure is the number of HH episodes in which the observed discharge score is the same as or higher than the expected discharge score, as determined in Step 4.

Step 7: Calculate the HHA-level discharge function percent. Divide the HHA's numerator count (Step 6) by its denominator count (Step 5) to obtain the HHA-level discharge function percent, then multiply by 100 to obtain a percent value.

Step 8: Round the percent value to two decimal places. If the digit in the third decimal place is 5 or greater, add 1 to the second decimal place, otherwise leave the second decimal place unchanged. Drop all digits following the second decimal place.

4. Measure Testing

4.1 Reportability

Reportability testing examines the total number and proportion of episodes that would have at least 20 eligible episodes for the Discharge Function Score measure in the reporting period. In CY 2021, 7,735 of a total of 9,764 HHAs (79.2%) met this threshold. This indicates high reportability and usability of the measure.

Exhibit 3. Publicly Reportable HHAs, CY 2021

Total Number of HHAs with ≥ 20 episodes	Percentage of HHAs with ≥ 20 episodes
7,735	79.2%

4.2 Variability

Variability testing summarizes the distribution of the facility-level final Discharge Function scores. In CY 2021, the mean final score among HHAs with at least 20 episodes was 58.0% (median: 61.6%, IQR: 49.7% - 70.0%). Final scores ranged from a minimum of 0.0% to a maximum of 100.0%. This wide variation indicates there is a performance gap in Discharge Function Scores across HHAs.

Exhibit 4. HHA-Level Distribution of Discharge Function Scores

N	Mean Score	Std dev.	Minimum	25th percentile	50th percentile	75th percentile	Maximum
7,735	58.0%	18.3%	0.0%	49.7%	61.6%	70.0%	100.0%

4.3 Reliability

The split-half reliability test examined agreement between two Discharge Function Scores for an HHA based on randomly-split, independent subsets of episodes in the same measurement period. Good agreement between the two measure scores calculated in this manner provides evidence that the measure is capturing an attribute of the HHA (quality of care) rather than the patient episodes (case-mix). For HHAs with at least 20 eligible episodes in CY 2021, each HHA's episodes were randomly divided into halves, thus ensuring that patient episodes were evenly distributed across the split-halves. Provider measure scores for each split-half sample were calculated. The Shrout-Fleiss intraclass correlation coefficient (ICC (2, 1)) was calculated between the split-half scores to measure reliability, applying the Spearman-Brown

correction.[17] The intraclass correlation coefficient for HHAs with more than 20 eligible episodes was 0.94, which indicates good reliability.[18]

4.4 Validity

This section reviews validity tests conducted to support the Discharge Function Score measure. Section 4.4.1 reports results that support the validity of measure scores. Section 4.4.2 describes analyses validating the imputation model results.

4.4.1 Measure Scores

To evaluate the validity of measure scores, convergent validity with other HH QRP measures, face validity, and risk adjustment model performance were assessed. The following subsections describe comparisons with other measures, webinars convened to gather expert, patient, and caregiver perspectives, and risk adjustment model calibration and fit analyses.

Convergent Validity

To evaluate convergent validity, the relationships between the Discharge Function Score measure and related HH QRP measures were examined. Using Spearman's rank correlation, the Discharge Function Score measure was compared to claims-based measure Discharge to Community (DTC) and to the assessment-based functional improvement measures (Improvement in Ambulation, Improvement in Bed Transfer, Improvement in Bathing, Improvement in Dyspnea, and Improvement in Oral Medication Management). The analysis used CY 2021 data from providers with at least 20 episodes. As shown in Exhibit 5, the Discharge Function Score measure was positively correlated with DTC (0.25) and each of the functional improvement measures: Improvement in Ambulation (0.25), Improvement in Bed Transfer (0.31), Improvement in Bathing (0.26), Improvement in Dyspnea (0.26), and Improvement in Oral Medication Management (0.23). All results were statistically significant ($p<0.05$). These results matched expectation. Higher functional status corresponds with higher likelihood of community discharge.[19] The other functional improvement measures are similarly positively correlated. We should expect this result given that these measures are also measuring patient function; however, there are two key differences between these measures and the

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

[18] Koo T.K. & Li M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine, 2016, 15(2), 155-163.

[19] Minor M, Jaywant A, Toglia J, Campo M, O'Dell MW. Discharge Rehabilitation Measures Predict Activity Limitations in Patients with Stroke Six Months after Inpatient Rehabilitation. Am J Phys Med Rehabil. 2021 Oct 20. doi: 10.1097/PHM.0000000000001908. Epub ahead of print. PMID: 34686630.

Discharge Function Score, which result in more modest positive correlations. First, the functional improvement measures measure whether the HHA improved patient function, while the Discharge Function Score measures whether patient function exceeds expectations at EOC. Second, the Discharge Function Score is a composite score of a spectrum of self-care and mobility function items, while the functional improvement measures focus on particular items.

Exhibit 5. Correlations between Discharge Function Score and Other Publicly Reported Measures

Measure	Spearman's Correlation	p-value
Discharge to Community–PAC HH QRP (CBE ID #3477)	0.25	<0.01
Improvement in Ambulation – Locomotion (CBE ID #0167)	0.25	<0.01
Improvement in Bed Transferring (CBE ID #0175)	0.31	<0.01
Improvement in Bathing (CBE ID #0174)	0.26	<0.01
Improvement in Dyspnea (CBE ID #0179)	0.26	<0.01
Improvement in Management of Oral Medications (CBE ID #0176)	0.23	<0.01

Face Validity

To assess face validity of the Discharge Function Score measure, two Technical Expert Panel (TEP) meetings (July 2021 and January 2022), as well as a Patient and Family Engagement Listening Session, were convened. TEP members showed strong support for the face validity of this measure. Though a vote was not taken at the meeting, the TEP agreed with the conceptual and operational definition of the measure. Panelists reviewed the validity analyses described herein and agreed they demonstrated measure validity.

The Patient and Family Engagement Listening Session demonstrated that the measure concept resonates with patients and caregivers. Participants' views of self-care and mobility were aligned with the functional domains captured by the measure, and they found them to be critical aspects of care. Participants emphasized the importance of measuring functional outcomes and were specifically interested in metrics that show how many patients discharged from particular HHAs made improvements in self-care and mobility.

Risk Adjustment Model Performance

The risk adjustment model is an ordinary least squares (OLS) linear regression. We assessed risk adjustment model calibration and fit using CY 2021 data. A well-calibrated model demonstrates good predictive ability to distinguish high-risk from low-risk patients. To assess risk adjustment model calibration, the ratios of observed-to-predicted discharge function score across eligible episodes by decile of predicted discharge function score (risk) were calculated. The average ratios of observed-to-predicted scores for each risk decile ranged from 0.98 to 1.01,

which suggested good calibration across the range of patients without evidence of concerning under- or over-estimation. Model fit was analyzed using adjusted R-squared to determine if the risk adjustment model can accurately predict discharge function while controlling for patient case-mix. The adjusted R-squared value was 0.49, which suggests good model discrimination.

4.4.2 Imputation Model

This section discusses the validity testing results of the imputation models used to estimate missing item scores. Validity testing included (1) assessments of model results and (2) calculation of bias and error of imputed item scores.

Model Results

To assess the validity of the imputation models, model fit, and face validity of model coefficients were evaluated. The C-statistic is a measure of model discrimination that determines the probability that predicting the outcome is better than chance. The C-statistic can range from 0.5 to 1, with 1 being perfect prediction and 0.5 being random chance. Using CY 2021 data, the C-statistic averaged 0.95 and ranged from 0.84 to 0.99 across the imputation models for each item at both admission and discharge (see Exhibit A-2). These results suggest good model discrimination across all imputation models.

The face validity of model results was assessed by reviewing model coefficients. For each item at both admission and discharge, imputation models produced sensible coefficients. Worse health conditions generally predicted lower item scores, as did prior functional status. Coefficients on related GG items were positively predictive and larger for GG items more closely related to the item being imputed (e.g., bed mobility items were generally more predictive for a bed mobility item imputation model than transfer or ambulation item imputation models).[20]

Bias and Mean Squared Error

A bootstrapping method was used to measure bias and mean squared error (MSE) in the imputation method. Bias measures the average amount by which the imputed value differs from the true value. Bias is signed, with a positive amount meaning that the imputed values were higher, on average, than the true values. MSE measures how far away the method is, on average from the truth. It is unsigned and can be positive even if bias is zero. The absolute size of bias is an inverse measure of accuracy, while the size of MSE is an inverse measure of the combination of precision and accuracy. The goal of the bootstrapping method was to determine how similar imputed values were to the true item score. This similarity could not be measured directly since

[20] Detailed model results are available upon request.

the true value of the measure score was unknown in the case of the individuals for whom imputation was necessary (imputation was needed precisely because the missing values prevented calculating the measure score for these individuals). Therefore, a bootstrapping strategy was implemented using the following steps to assess the accuracy of the statistical imputation method:

Step 1: Identified observations from the original sample with no NAs recorded across all items needed for measure calculation.

Step 2: Generated a bootstrap sample that draws from the no-NA observations until there were as many observations in the bootstrap sample as the original sample. A stratified random sampling algorithm was used. The first stratum of each bootstrap sample consisted of no-NA observations. This stratum had the same number of observations as there were no-NA observations in the original data. This stratum of the bootstrap sample was filled by simple random sampling from the no-NA observations.

To fill the bootstrap sample observations corresponding to the observations from the original data having NAs, it was not possible to use simple random sampling. This is because the distribution of clinical and function characteristics was different between observations with and without NAs. Therefore, the sampling to fill the bootstrap sample for these observations was done using a stratification method which matched observations with NA to similar observations without NA.

Therefore, ten additional strata were filled corresponding to the observations from the original data with NAs. These strata were defined by the deciles of a predicted score estimated, as described in Section 3.5. Bootstrap observations corresponding to the observations with NAs were chosen by simple random sampling within each of these strata.

Step 3: Created two copies of this sample.

a) One copy served as the gold standard source of truth because all observations in the bootstrap sample were sampled from no-NA observations.

b) In the other copy, NAs were imposed on some of the GG items. This was done in a way which preserved both the pattern of NAs within the data and the pattern of clinical characteristics among NA observations. NAs were imposed by randomly selecting observations from the original data which i) had NAs and ii) were in the same stratum (see Step 2) as the corresponding target observation in the second copy. The GG items which were missing in the sampled observation were made missing in the target observation.

Step 4: In the second copy produced in Step 3b, imputed values for the NAs imposed onto the bootstrap sample were generated. For comparison, applied "recode to 1" method and calculated resulting measure scores.

Step 5: Calculated bias and mean-squared error of the imputation method by comparing observation by observation to the measure scores produced from the gold standard copy (Step 3a).

Step 6: Repeated Steps 2-5 many times. Reported average bias/mean-squared error across iterations/bootstrap replications.

Bias and MSE were compared between statistical imputation and the current method for in-use measures, which recodes all NAs to 1. Using this bootstrapping method, statistical imputation resulted in lower levels of bias (-0.19 at admission; -0.13 at discharge) and MSE (1.64 at admission; 1.10 at discharge) compared to the bias (-0.28 at admission; -0.53 at discharge) and MSE (1.84 at admission; 10.35 at discharge) produced from the recode approach, which supports the validity of the imputation method.

Appendix A

Exhibit A-1. Discharge Function Risk Adjustment: Linear Regression Model Results, CY 2021

Covariate	Number of Episodes	Percent of Episodes	Average Observed Score	Estimate	P-value
Age Group < 35 years	48,837	1%	52.37	-0.30	0.00
Age Group 35 - 44 years	66,666	1%	53.22	-0.28	0.00
Age Group 45 - 54 years	153,444	3%	53.88	-0.30	0.00
Age Group 55 - 64 years	413,644	9%	54.26	-0.23	0.00
Age Group 75 - 84 years	1,422,575	32%	53.88	-0.09	0.00
Age Group 85 - 90 years	670,543	15%	52.35	-0.40	0.00
Age Group > 90 years	391,088	9%	50.27	-1.08	0.00
Admission Mobility - continuous form	.	.	.	1.01	0.00
Admission Mobility - squared form	.	.	.	-0.01	0.00
Prior Surgery (having a diagnosis ICD-10 code between Z40 and Z53)	1,125,431	25%	56.45	0.96	0.00
Prior Functioning: Self Care Dependent	194,957	4%	33.86	-4.25	0.00
Prior Functioning: Self Care Some Help	1,828,761	41%	51.20	-0.94	0.00
Prior Functioning: Indoor Mobility (Ambulation) - Dependent	221,151	5%	34.90	-2.24	0.00
Prior Functioning: Indoor Mobility (Ambulation) - Some Help	1,653,018	37%	51.08	-1.35	0.00
Prior Functioning: Stairs - Dependent	879,198	20%	47.43	-0.43	0.00
Prior Functioning: Stairs - Some Help	1,553,571	35%	52.69	0.04	0.00
Prior Functioning: Functional Cognition - Dependent	275,834	6%	40.30	-0.94	0.00
Indicator: wheeler	184,939	4%	35.35	-7.00	0.00
Prior Mobility Device Use: Walker	2,198,354	49%	52.61	-0.28	0.00
Prior Mobility Device Use: Wheelchair	639,180	14%	44.24	-1.96	0.00
Prior Mobility Device Use: Mechanical Lift	55,217	1%	34.74	-3.37	0.00
Prior Mobility Device Use: Orthotics/Prosthetics	43,902	1%	51.65	0.88	0.00
Stage 2 Pressure Ulcer	97,972	2%	43.94	-1.95	0.00
Stage 3, 4 or Unstageable Pressure Ulcer/Injury	60,864	1%	41.73	-2.48	0.00
Cognitive Function: Moderately Impaired	1,790,216	40%	51.95	-0.54	0.00
Cognitive Function: Severely Impaired	183,839	4%	40.26	-3.04	0.00
Bladder Incontinence: Admission - Incontinent	2,103,618	47%	51.25	-0.55	0.00

Covariate	Number of Episodes	Percent of Episodes	Average Observed Score	Estimate	P-value
Bladder Incontinence: Catheter	138,819	3%	48.97	-1.48	0.00
Bowel Continence: Always incontinent	206,970	5%	42.24	-2.57	0.00
Bowel Continence: Incontinent Less than Daily	499,743	11%	47.77	-1.00	0.00
Availability of Assistance: Around the Clock	3,280,019	73%	53.00	-0.45	0.00
Availability of Assistance: Regular Daytime	198,028	4%	53.89	-0.35	0.00
Availability of Assistance: Regular Nighttime	164,415	4%	55.44	0.04	0.05
Living Arrangements: Live Alone	1,064,998	24%	55.94	0.59	0.00
Living Arrangements: Congregate Setting	429,428	10%	49.00	-0.05	0.00
Start of Care-Further visits planned; Not Discharged from facility in past 14 days	1,566,688	35%	52.02	-0.87	0.00
Resumption of care (after inpatient stay)	349,011	8%	51.60	-0.67	0.00
Low Body Mass Index	195,959	4%	51.62	-0.50	0.00
High Body Mass Index (BMI > 50)	67,278	2%	52.80	-0.62	0.00
Risk for hospitalization: History of Falls	1,731,061	39%	53.04	0.29	0.00
Risk for hospitalization: Multiple hospitalizations	1,328,512	30%	53.08	-0.02	0.03
Risk for hospitalization: Multiple ER visits	1,280,586	29%	53.06	0.19	0.00
Risk for hospitalization: Decline in status	1,542,176	34%	51.98	0.07	0.00
Risk for hospitalization: Non-compliance	1,973,298	44%	52.85	0.22	0.00
Risk for hospitalization: Currently taking 5 or more medications	4,201,022	94%	53.70	0.26	0.00
Confusion: Moderate	1,700,358	38%	52.98	-0.09	0.00
Confusion: Severe	499,282	11%	46.10	-1.11	0.00
Vision: Partial impairment	1,218,545	27%	51.41	-0.42	0.00
Vision: Severe impairment	52,663	1%	47.26	-1.40	0.00
Needs oral medication management	3,217,344	72%	53.09	0.40	0.00
Needs injectable medication management	607,917	14%	52.91	-0.18	0.00
Supervision and safety: Non-agency caregiver(s) currently provide assistance	1,850,424	41%	53.10	-0.29	0.00
Supervision and safety: Non-agency caregiver(s) need training/supportive services to provide assistance	1,623,234	36%	52.72	-0.02	0.02

Covariate	Number of Episodes	Percent of Episodes	Average Observed Score	Estimate	P-value
Supervision and safety: Non-agency caregiver(s) are not likely to provide assistance	106,433	2%	54.49	-0.06	0.02
Supervision and safety: Assistance needed, but no non-agency caregiver(s) available	81,931	2%	54.97	-0.38	0.00
Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock (HCC2)	82,638	2%	53.35	-0.03	0.23
Metastatic Cancer and Acute Leukemia (HCC8)	60,428	1%	52.96	-1.84	0.00
Lung and Other Severe Cancers (HCC9)	69,066	2%	54.40	-0.97	0.00
Breast, Prostate, and Other Cancers and Tumors (HCC12)	71,573	2%	54.63	-0.12	0.00
Diabetes without Complication (HCC19)	504,358	11%	54.09	0.02	0.15
Protein-Calorie Malnutrition (HCC21)	58,578	1%	51.95	-0.40	0.00
Intestinal Obstruction/Perforation (HCC33)	28,341	1%	54.88	0.51	0.00
Inflammatory Bowel Disease (HCC35)	18,420	0%	55.68	0.75	0.00
Bone/Joint/Muscle Infections/Necrosis (HCC39)	57,459	1%	54.24	-0.19	0.00
Rheumatoid Arthritis and Inflammatory Connective Tissue Disease (HCC40)	131,039	3%	54.57	0.12	0.00
Dementia With Complications (HCC51)	80,818	2%	45.80	-1.86	0.00
Dementia Without Complication (HCC52)	384,481	9%	47.55	-1.31	0.00
Quadriplegia (HCC70)	8,789	0%	31.56	-4.21	0.00
Paraplegia (HCC71)	14,137	0%	42.67	1.11	0.00
Spinal Cord Disorders/Injuries (HCC72)	18,906	0%	50.66	-0.08	0.11
Amyotrophic Lateral Sclerosis and Other Motor Neuron Disease (HCC73)	5,691	0%	37.55	-6.84	0.00
Cerebral Palsy (HCC74)	15,123	0%	38.20	-2.22	0.00
Muscular Dystrophy (HCC76)	3,499	0%	44.40	-1.75	0.00
Multiple Sclerosis (HCC77)	36,244	1%	47.91	-0.66	0.00
Parkinson's and Huntington's Diseases (HCC78)	137,681	3%	48.90	-1.45	0.00
Seizure Disorders and Convulsions (HCC79)	107,309	2%	49.95	-0.53	0.00
Respirator Dependence/Tracheostomy Status (HCC82)	6,168	0%	49.25	-1.90	0.00
Congestive Heart Failure (HCC85)	874,066	20%	53.27	-0.30	0.00
Acute Myocardial Infarction (HCC86)	74,614	2%	55.29	0.12	0.00
Specified Heart Arrhythmias (HCC96)	649,912	15%	54.13	0.22	0.00
Ischemic or Unspecified Stroke (HCC100)	12,919	0%	50.56	-1.20	0.00
Hemiplegia/Hemiparesis (HCC103)	201,428	5%	48.46	-1.51	0.00

Covariate	Number of Episodes	Percent of Episodes	Average Observed Score	Estimate	P-value
Atherosclerosis of the Extremities with Ulceration or Gangrene (HCC106)	13,584	0%	51.46	-1.20	0.00
Pneumococcal Pneumonia, Empyema, Lung Abscess (HCC115)	12,974	0%	54.38	0.03	0.64
Exudative Macular Degeneration (HCC124)	1,743	0%	54.70	0.79	0.00
Dialysis Status (HCC134)	16,955	0%	52.17	-0.96	0.00
Chronic Kidney Disease, Stage 5 (HCC136)	88,725	2%	52.32	-0.42	0.00
Chronic Kidney Disease, Moderate (Stage 3) (HCC138)	353,823	8%	53.82	0.19	0.00
Hip Fracture/Dislocation (HCC170)	5,168	0%	53.11	-0.02	0.80
Complications of Specified Implanted Device or Graft (HCC176)	25,129	1%	54.85	0.32	0.00
Amputation Status, Lower Limb/Amputation Complications (HCC189)	28,403	1%	52.11	0.34	0.00
Intercept	.	.	.	33.74	0.00

Exhibit A-2. C-Statistics for Imputation Models across GG Items at Admission and Discharge, CY 2021

Item	Description	Assessment Timing	C Statistic
GG0130A	Eating	Admission	0.85
		Discharge	0.94
GG0130B	Oral Hygiene	Admission	0.87
		Discharge	0.97
GG0130C	Toileting Hygiene	Admission	0.92
		Discharge	0.97
GG0170A	Roll left/right	Admission	0.93
		Discharge	0.98
GG0170C	Lying to sit – bed	Admission	0.97
		Discharge	0.99
GG0170D	Sit to stand	Admission	0.97
		Discharge	0.99
GG0170E	Chair to bed trans.	Admission	0.98
		Discharge	0.99
GG0170F	Toilet trans.	Admission	0.97
		Discharge	0.98
GG0170I	Walk 10'	Admission	0.95
		Discharge	0.98
GG0170J	Walk 50'	Admission	0.97
		Discharge	0.98
GG0170R	Wheel 50'	Admission	0.84
		Discharge	0.86