Discharge Function Score for Home Health (HH)

Technical Report

Updated February 2024

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 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 positively 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 cross-setting Discharge Function Score (DC Function) measure determines how successful each HHA is at achieving or exceeding an expected level of functional ability for its patients at discharge. An expectation for discharge function score is built for each HHA quality episode by accounting for patient characteristics that impact their functional status. The final cross-setting DC Function for a given HHA is the proportion of that HHA's quality episodes where a patient's observed discharge function score meets or exceeds their expected discharge function score. HHAs with low scores indicate that they are not achieving the functional gains at discharge that are expected based upon patient characteristics and patient status at start of care (SOC) or resumption of care (ROC) for a larger share of their patients. The measure provides information to HHAs that has the potential to hold providers accountable for functional

[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.

outcomes and encourages them to improve the quality of care they deliver. This measure also promotes patient wellness, encourages adequate nursing and therapy services to help prevent adverse outcomes (e.g., potentially preventable hospitalization) and increases the transparency of quality of care in the HH setting. DC Function 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 activity scores including those coded with activity not attempted codes.

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 technical measure specifications for DC Function. Section 2 provides an overview of the measure and is a high-level summary of the key features of the measure. Section 3 describes the methodology used to construct DC Function including its data sources, study population, measure outcome, and steps for calculating the final measure score. Section 4 discusses DC Function 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 activity scores.

2. Overview

This section provides an overview of basic descriptive information on the DC Function 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 (DC Function)

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

DC Function calculates the percent of HH patients who achieve or exceed an expected discharge function score. Functional status is measured through Section GG of OASIS assessments, which evaluates a patient's performance of daily activities related to self-care (GG0130) and mobility (GG0170). OASIS-derived coefficients from a risk adjustment model controlling for SOC/ROC function score, age, and patient clinical characteristics are used to determine an expected discharge function score for each HH quality episode. The provider score is calculated as the following proportion:

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

Observed Discharge Function Score is calculated using Section GG data from OASIS assessments at discharge (M0100 Reason for Assessment = 9).

Expected Discharge Function Score estimates an expected discharge function score by using risk-adjusted OASIS data from SOC/ROC (M0100 Reason for Assessment = 1 or 3).

3. Measure Specifications

3.1 Measure Time Period

This measure is calculated using 12 months (four quarters) of data. All HH quality episodes with a discharge date that falls within this reporting 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 currently collected on all Medicare fee-for-service (FFS), Medicare Advantage, Medicaid FFS, and Medicaid managed care patients who receive services from an HHA with the following exceptions: patients under the age of 18, patients receiving maternity services, patients receiving only personal care, housekeeping services, or chore services. There will be no additional data collection or submission burden for HHAs as agencies are currently collecting data for the OASIS items used in the measure calculation.

3.3 Denominator

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

3.3.1 HH Quality Episode Construction

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

3.3.2 Eligible Quality Episodes

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

- The quality episode is defined as an incomplete stay by meeting one of the following criteria:
    - Quality episodes that end in a transfer (M0100 reason for assessment = 6 or 7) during the reporting period
    - Quality episodes that end with Death at Home (M0100 reason for assessment = 8); and
    - Quality episodes lasting 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. The quality episode is for a patient considered to be non-responsive, in which the primary diagnosis (M1021) or other diagnoses (M1023) indicates that the patient has a diagnosis of coma, persistent vegetative state, complete tetraplegia, locked-in state, severe anoxic brain damage, cerebral edema, or compression of brain and in which the patient's cognitive functioning (M1700) is totally dependent due to disturbances such as constant disorientation, coma, persistent vegetative state, or delirium.
    - Rationale: These patients are excluded because they may have limited or less predictable functional abilities.
- Patient is discharged to hospice (home or institutional facility)
    - Rationale: Patient priorities may change during the HH quality episode for a patient discharged to hospice.

3.4 Numerator

The numerator is the number of quality episodes during the reporting period in which the observed discharge function score (Section 3.4.1) for select GG function activities is equal to or greater than the 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 activities at discharge. Section GG of each PAC assessment instrument includes standardized patient assessment data elements that measure functional status. DC Function measure focuses on GG activities that are currently available across all PAC settings (Exhibit 1).

Exhibit 1. Cross-Setting Function Item Set

Item/Activity	Description
GG0130A	Eating
GG0130B	Oral Hygiene
GG0130C	Toileting Hygiene
GG0170A	Roll Left and Right
GG0170C	Lying to Sitting on Side of Bed
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/activities are reported in Exhibit 2.

Exhibit 2. GG Items Response

Category	GG Items Response	Response Description
Patient Functional Status Assessed	06	Independent
	05	Setup or clean-up assistance
	04	Supervision or touching assistance
	03	Partial/moderate assistance
	02	Substantial/maximal assistance
	01	Dependent
Activity Not Attempted (ANA) codes	07	Patient refused
	09	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 activity is between 01 and 06, then use code as the score for that activity.

Step 2: If code for an activity is 07, 09, 10, 88, dashed (-), skipped (^), or missing, then use statistical imputation to estimate the activity score for that activity (see Section 3.5).

Step 3: Sum scores across all activities to calculate the total observed discharge function score.

Step 4: Round the observed discharge function score to the fourth decimal place.

Different locomotion activities are used if the patient is wheelchair-bound than for the remaining patients:

Use 2 * Wheel 50 Feet with 2 Turns (GG0170R) score to calculate the total observed discharge function score for quality episodes where (i) Walk 10 Feet (GG0170I) has an activity not attempted (ANA) code at both SOC/ROC and discharge and (ii) either Wheel 50 Feet with 2 Turns (GG0170R) has a code between 01 and 06 at either SOC/ROC or discharge. The remaining quality 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 activities 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 quality episode using SOC/ROC OASIS data. Risk adjustment controls for patient characteristics such as SOC/ROC function score, age, and clinical conditions. Refer to Section 3.6 for details on risk adjustment. For consistent comparison against the observed discharge function score, the expected discharge function score is also rounded to the fourth decimal place.

3.5 Statistical Imputation

When an activity 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, activity scores are estimated through a process known as statistical imputation. On average, patients who are coded as NA on a GG activity at SOC/ROC tend to score higher at discharge (if assessed) than patients who are coded as dependent at SOC/ROC. 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 01 to 06, based on a patient's clinical characteristics determined by the OASIS assessment including codes assigned on other GG activities.

A separate statistical imputation model is constructed for each GG activity used in DC Function (Section 3.4.1) at SOC/ROC and at discharge. Imputation models include the predictors used in risk adjustment (Section 3.6.2) and covariates for scores on other GG activities (Step 3 below). Notably, imputation models use all GG activities available in HH to estimate missing scores for the subset of GG activities used for the DC Function numerator. Detailed imputation model results are available in the downloads section in the DC-Function-Imputation-Appendix-HH document (https://www.cms.gov/medicare/quality/home-health/home-health-quality-measures).

The following steps are used to generate imputed activity scores for quality episodes with NA codes. Note that these steps first describe imputing a single activity at SOC/ROC and then describe the relevant modifications for imputing that activity at discharge and for the other activities.

Step 1: Start with Eating (GG0130A). Identify eligible quality episodes where the activity score is not missing (i.e., had a score 01-06) at SOC/ROC. These scores are used as the outcome (i.e., left-hand-side variable) of the SOC/ROC imputation model for GG0130A.

Step 2: For each quality episode, determine whether to use walking or wheelchair activities in the imputation model.

a) If Walk 10 Feet (GG0170I) has an ANA code at both SOC/ROC and discharge and either Wheel 50 Feet with 2 Turns (GG0170R) or Wheel 150 Feet (GG0170S) has a code between 01 and 06, then use wheelchair activities.

b) Otherwise, use walking activities.

Step 3: Create variables for the imputation model reflecting how each activity except Eating (GG0130A) was scored at SOC/ROC. GG activity 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 activity had an ANA code, dash, or missing value (g^*), and the third is an indicator variable taking a value of 1 if the activity was skipped (g^{**}).

$$Function\ activities : G \in \{g_2, \dots, g_{10}\} \tag{1}$$

$$g' = [g, g^*, g^{**}] \tag{2}$$

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

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

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

$$Function\ activities\ with\ NA\ indicators : G' \in \{g'_2, \dots, g'_{10}\} \tag{3}$$

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 activities with NA indicators (G') constructed in Step 3.

$$Clinical\ items := C \in \{c_1, \dots, c_k\} \tag{4}$$

$$Function\ activities\ with\ NA\ indicators : G' \in \{g'_2, \dots, g'_{10}\} \tag{5}$$

The model we estimate for g_1, GG0130A, is

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

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

The latent variable, z_i, is interpreted as patient i's underlying degree of independence on assessment activity 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 activity, 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 activity. 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 (rounded to four decimal places) as

$$\begin{aligned} \widehat{g_i} = \Pr(z_i \leq \alpha_1) + 2 * \Pr(\alpha_1 < z_i \leq \alpha_2) + 3 * \Pr(\alpha_2 < z_i \leq \alpha_3) + \\ 4 * \Pr(\alpha_3 < z_i \leq \alpha_4) + 5 * \Pr(\alpha_4 < z_i \leq \alpha_5) + 6 * \Pr(z_i > \alpha_5) \end{aligned} \tag{8}$$

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

Step 6: Repeat Steps 1 - 5 for each GG activity included in the observed discharge function score (Section 3.4.1), as above replacing the Eating (GG0130A) activity with each successive GG activity 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 quality 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 quality episode that controls for SOC/ROC functional status, age, and clinical characteristics. This ensures that each quality episode is measured against an expectation that is calibrated to the patient's individual circumstances at SOC/ROC 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 quality episode, incorporating imputed activity scores when NA codes are encountered. The risk adjustment model is run on all HHA quality episodes to determine the model intercept (β_0) and risk adjustor coefficients ($\beta_1, \ \dots, \beta_n$). Expected discharge function scores are calculated by applying the regression equation to each HHA quality episode at SOC/ROC.

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

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 measures. Information on the covariates were obtained from the SOC/ROC OASIS data.

Age Category

Age was calculated as of the SOC/ROC date (M0030/M0032) of the HH quality episode using the patient's date of birth (M0066).

SOC/ROC Function Score

Sum of SOC/ROC scores for function activities included in the discharge score (Section 3.6.1), which can range from 10-60, with a higher score indicating greater independence. NAs in the SOC/ROC activity scores are treated the same way as NAs in the discharge activity scores, with NAs replaced with imputed scores (Steps 1-2 in Section 3.4.1). The walking and wheelchair activities are used in the same manner as for the discharge score (Step 3 in Section 3.6.1). SOC/ROC function 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 quality episode.

Pressure Ulcers

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

Cognitive Function

These covariates capture the patient's cognitive function by assessing whether the patient's mental status at SOC/ROC 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.

SOC/ROC 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.

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.

3.7 Measure Calculation

DC Function is the proportion of HH quality episodes in which the observed discharge function score is equal to or greater than an expected discharge function score . A higher score indicates better performance in functional outcomes. For each HH quality episode, observed discharge function score (Section 3.4.1) and expected discharge function score (Section 3.4.2) are determined. For each HHA, DC Function is the proportion of quality episodes where the observed discharge function score is greater than or equal to the expected discharge function score.

3.7.1 Steps Used in Calculation

Step1: Calculate the observed discharge function score (M0100 reason for assessment = 9) as described in Section 3.4.1, incorporating imputed activity scores (Section 3.5).

Step 2: Identify excluded HH quality episodes using the criteria mentioned in Section 3.3.

Step 3: Calculate the expected discharge function score (M0100 reason for assessment = 1 or 3). For each HH quality 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 quality episode which does not meet the exclusion criteria, compare each patient'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 function score is equal to or greater than the expected discharge function score.
- Observed discharge function score is lower than the expected discharge function score.

Step 5: Determine the denominator count. Determine the total number of HH quality episodes with an OASIS discharge date in the measure reporting 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 quality episodes in which the observed discharge function score (rounded to four decimal places) is the equal to or greater than the expected discharge function score (rounded to four decimal places).

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 the digits following the second decimal place.

4. Measure Testing

4.1 Reportability

Reportability testing examines the total number and proportion of quality episodes that would have at least 20 eligible quality episodes for the Discharge Function measure in the reporting period. In CY 2022, 81.1% (n=8,168) of total HHAs (n=10,069) met this threshold. This indicates high reportability and usability of the measure.

Exhibit 3. Publicly Reportable HHAs, CY 2022

Total Number of HHAs	Percentage of HHAs with ≥ 20 episodes
10,069	81.1%

4.2 Variability

Variability testing summarizes the distribution of the agency-level final DC Function. In CY 2022, the mean final score among HHAs with at least 20 quality episodes was 57.5% (median: 61.4%, IQR: 48.7% - 69.8%). 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 DC Function across HHAs.

Exhibit 4. HHA-Level Distribution of DC Function

N	Mean Score	Std dev.	Minimum	25th percentile	50th percentile	75th percentile	Maximum
8,168	57.5%	18.7%	0.0%	48.7%	61.4%	69.8%	100.0%

4.3 Reliability

The split-half reliability test examined agreement between two Discharge Function measure scores for a HHA based on randomly-split, independent subsets of quality episodes 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 HHA (quality of care) rather than the patient episodes (case-mix). For HHAs with at least 20 eligible quality episodes in CY 2022, each HHA's quality episodes were randomly divided into halves, thus ensuring that patient quality 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.[15] The intraclass correlation coefficient for HHAs with more than 20 eligible quality episodes was 0.98, which indicates good reliability.[16]

4.4 Validity

This section reviews validity tests conducted to support DC Function. 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 DC Function and related HH QRP measures were examined. Using Spearman's rank correlation, DC Function was compared to claims-based measure Discharge to Community (DTC) and to the assessment-based functional improvement measures (Improvement in Ambulation/Locomotion, Improvement in Bed Transfer, Improvement in Bathing, Improvement in Dyspnea, and Improvement in Management of Oral Medications). The analysis used CY 2022 data from providers with at least 20 quality episodes. As shown in Exhibit, DC Function was positively correlated with DTC (0.27) and each of the functional improvement measures: Improvement in Ambulation/Locomotion (0.28), Improvement in Bed Transfer (0.38), Improvement in Bathing (0.26), Improvement in Dyspnea (0.33), and Improvement in Management of Oral Medication (0.24). All results were statistically significant ($p<0.05$). These results matched expectations. Higher functional status corresponds with higher likelihood of community discharge.[17] 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 DC Function, which result in more modest positive correlations. First, the functional

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

[16] 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.

[17] 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.

improvement measures measure whether the HHA improved patient function, while DC Function whether patient function exceeds expectations at discharge. Second, DC Function is a composite score of a spectrum of self-care and mobility function activities, while the functional improvement measures each focus on one specific functional item.

Exhibit 5. Correlations between DC Function and Other Publicly Reported Measures

Measure	Spearman's Correlation	p-value
Discharge to Community–PAC HH QRP (NQF #3477)	0.27	<0.01
Improvement in Ambulation – Locomotion (NQF #0167)	0.28	<0.01
Improvement in Bed Transferring (NQF #0175)	0.38	<0.01
Improvement in Bathing (NQF #0174)	0.26	<0.01
Improvement in Dyspnea (NQF #0179)	0.33	<0.01
Improvement in Management of Oral Medications (NQF #0176)	0.24	<0.01

Face Validity

To assess face validity of DC Function, 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 2022 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-expected discharge function score across eligible quality episodes by decile of expected discharge function score (risk) were calculated. The average ratios of observed-to-expected scores for each risk decile ranged from 0.95 to 1.00, 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 how the validity of the imputation models used to estimate missing activity scores was determined: (1) review of model results, (2) calculation of bias and error of imputed activity scores, and (3) a comparison of discharge function between episodes with assessed and imputed SOC/ROC 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 2022 data, the C-statistic averaged 0.95 and ranged from 0.83 to 0.99 across the imputation models for each item at both SOC/ROC 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 activity at both SOC/ROC and discharge, imputation models produced sensible coefficients. Worse health conditions generally predicted lower item scores, as did prior functional status. Scores on related GG activities were positively predictive, and GG activities generally were more predictive the more similar were the functions being measured (e.g., bed mobility items were generally more predictive of other bed mobility activities than, for example, were transfer or ambulation items).[18]

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 were 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

[18] Detailed model results are available upon request.

how similar imputed values were to the true activity score. This similarity could not be measured directly since 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 strata 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 activities. 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 activities 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 (-0.22 at SOC/ROC; -0.15 at discharge) and MSE (1.44 at SOC/ROC; 1.23 at discharge) compared to the bias (-0.54 at SOC/ROC; -0.70 at discharge) and MSE (4.60 at SOC/ROC; 13.30 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 2022

Covariate	Number of Episodes	Percent of Episodes	Average Observed Score	Estimate	P-value
Age Group < 35 years	51,723	1%	51.92	-0.34	0.00
Age Group 35 - 44 years	78,038	2%	52.92	-0.27	0.00
Age Group 45 - 54 years	180,105	3%	53.59	-0.31	0.00
Age Group 55 - 64 years	481,710	9%	54.01	-0.25	0.00
Age Group 75 - 84 years	1,642,696	32%	53.66	-0.10	0.00
Age Group 85 - 90 years	779,771	15%	52.14	-0.41	0.00
Age Group > 90 years	454,030	9%	50.06	-1.08	0.00
Admission Mobility - continuous form	-	-	-	0.99	0.00
Admission Mobility - squared form	-	-	-	-0.01	0.00
Prior Surgery (having a diagnosis ICD-10 code between Z40 and Z53)	1,250,249	24%	56.16	0.93	0.00
Prior Functioning: Self Care Dependent	236,647	5%	33.79	-4.27	0.00
Prior Functioning: Self Care Some Help	2,172,563	42%	51.06	-0.96	0.00
Prior Functioning: Indoor Mobility (Ambulation) - Dependent	268,697	5%	34.86	-2.19	0.00
Prior Functioning: Indoor Mobility (Ambulation) - Some Help	1,967,126	38%	50.95	-1.36	0.00
Prior Functioning: Stairs - Dependent	1,033,408	20%	47.14	-0.42	0.00
Prior Functioning: Stairs - Some Help	1,838,618	36%	52.53	0.04	0.00
Prior Functioning: Functional Cognition - Dependent	330,714	6%	40.16	-0.95	0.00
Indicator: wheeler	224,822	4%	35.28	-7.02	0.00
Prior Mobility Device Use: Walker	2,570,686	50%	52.42	-0.29	0.00
Prior Mobility Device Use: Wheelchair	766,920	15%	44.09	-1.93	0.00
Prior Mobility Device Use: Mechanical Lift	66,202	1%	34.34	-3.40	0.00
Prior Mobility Device Use: Orthotics/Prosthetics	51,679	1%	51.38	0.87	0.00
Stage 2 Pressure Ulcer	120,173	2%	43.78	-1.87	0.00
Stage 3, 4 or Unstageable Pressure Ulcer/Injury	77,872	2%	41.53	-2.37	0.00
Cognitive Function: Moderately Impaired	2,115,320	41%	51.73	-0.56	0.00
Cognitive Function: Severely Impaired	213,716	4%	40.03	-3.07	0.00
Bladder Incontinence: Admission - Incontinent	2,483,654	48%	51.04	-0.60	0.00

Covariate	Number of Episodes	Percent of Episodes	Average Observed Score	Estimate	P-value
Bladder Incontinence: Catheter	168,993	3%	48.21	-1.69	0.00
Bowel Continence: Always incontinent	246,767	5%	41.98	-2.57	0.00
Bowel Continence: Incontinent Less than Daily	597,610	12%	47.55	-1.02	0.00
Availability of Assistance: Around the Clock	3,787,091	73%	52.75	-0.46	0.00
Availability of Assistance: Regular Daytime	231,936	4%	53.65	-0.37	0.00
Availability of Assistance: Regular Nightime	189,220	4%	55.17	0.00	0.83
Living Arrangements: Live Alone	1,233,085	24%	55.74	0.59	0.00
Living Arrangements: Congregate Setting	501,151	10%	48.87	-0.05	0.00
Start of Care-Further visits planned; Not Discharged from facility in past 14 days	1,849,153	36%	51.77	-0.96	0.00
Resumption of Care (after inpatient stay)	424,446	8%	51.33	-0.75	0.00
Low Body Mass Index	227,600	4%	51.42	-0.48	0.00
High Body Mass Index (BMI > 50)	80,102	2%	52.52	-0.71	0.00
Risk for hospitalization: History of Falls	2,013,524	39%	52.86	0.32	0.00
Risk for hospitalization: Multiple hospitalizations	1,546,716	30%	52.90	0.00	0.75
Risk for hospitalization: Multiple ER visits	1,489,939	29%	52.89	0.22	0.00
Risk for hospitalization: Decline in status	1,785,430	35%	51.78	0.10	0.00
Risk for hospitalization: Non-compliance	2,293,892	44%	52.64	0.23	0.00
Risk for hospitalization: Currently taking 5 or more medications	4,855,991	94%	53.46	0.25	0.00
Confusion: Moderate	2,005,411	39%	52.75	-0.11	0.00
Confusion: Severe	585,871	11%	45.93	-1.10	0.00
Vision: Partial impairment	1,453,205	28%	51.18	-0.45	0.00
Vision: Severe impairment	62,536	1%	47.01	-1.50	0.00
Needs oral medication management	3,690,407	71%	52.84	0.41	0.00
Supervision and safety: Non-agency caregiver(s) currently provide assistance	2,130,624	41%	52.86	-0.31	0.00
Supervision and safety: Non-agency caregiver(s) need training/supportive services to provide assistance	1,898,462	37%	52.47	-0.03	0.00

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	125,222	2%	54.30	-0.05	0.02
Supervision and safety: Assistance needed, but no non-agency caregiver(s) available	97,823	2%	54.77	-0.40	0.00
Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock (HCC2)	94,800	2%	53.14	-0.04	0.10
Metastatic Cancer and Acute Leukemia (HCC8)	69,070	1%	52.94	-1.77	0.00
Lung and Other Severe Cancers (HCC9)	79,099	2%	54.34	-0.90	0.00
Breast, Prostate, and Other Cancers and Tumors (HCC12)	81,945	2%	54.42	-0.10	0.00
Diabetes without Complication (HCC19)	584,909	11%	53.86	-0.01	0.39
Protein-Calorie Malnutrition (HCC21)	67,213	1%	51.83	-0.36	0.00
Intestinal Obstruction/Perforation (HCC33)	32,342	1%	54.76	0.55	0.00
Inflammatory Bowel Disease (HCC35)	21,248	0%	55.57	0.78	0.00
Bone/Joint/Muscle Infections/Necrosis (HCC39)	68,736	1%	54.01	-0.17	0.00
Rheumatoid Arthritis and Inflammatory Connective Tissue Disease (HCC40)	151,139	3%	54.31	0.11	0.00
Dementia With Complications (HCC51)	94,075	2%	45.65	-1.88	0.00
Dementia Without Complication (HCC52)	449,546	9%	47.37	-1.34	0.00
Quadriplegia (HCC70)	11,166	0%	31.06	-4.39	0.00
Paraplegia (HCC71)	18,067	0%	42.28	1.10	0.00
Spinal Cord Disorders/Injuries (HCC72)	22,017	0%	50.23	-0.10	0.04
Amyotrophic Lateral Sclerosis and Other Motor Neuron Disease (HCC73)	6,793	0%	37.10	-7.06	0.00
Cerebral Palsy (HCC74)	18,251	0%	37.82	-2.40	0.00
Muscular Dystrophy (HCC76)	4,153	0%	43.89	-1.90	0.00
Multiple Sclerosis (HCC77)	43,043	1%	47.24	-0.81	0.00
Parkinson's and Huntington's Diseases (HCC78)	161,994	3%	48.67	-1.49	0.00
Seizure Disorders and Convulsions (HCC79)	125,786	2%	49.70	-0.56	0.00
Respirator Dependence/Tracheostomy Status (HCC82)	7,469	0%	48.67	-1.92	0.00
Congestive Heart Failure (HCC85)	1,018,521	20%	53.08	-0.32	0.00
Acute Myocardial Infarction (HCC86)	85,137	2%	55.16	0.14	0.00
Specified Heart Arrhythmias (HCC96)	749,228	14%	53.94	0.23	0.00
Ischemic or Unspecified Stroke (HCC100)	15,421	0%	50.31	-1.19	0.00
Hemiplegia/Hemiparesis (HCC103)	239,592	5%	48.20	-1.55	0.00

Covariate	Number of Episodes	Percent of Episodes	Average Observed Score	Estimate	P-value
Atherosclerosis of the Extremities with Ulceration or Gangrene (HCC106)	16,908	0%	51.51	-1.05	0.00
Pneumococcal Pneumonia, Empyema, Lung Abscess (HCC115)	15,116	0%	54.17	0.03	0.61
Exudative Macular Degeneration (HCC124)	1,994	0%	54.64	0.89	0.00
Dialysis Status (HCC134)	19,851	0%	52.08	-0.93	0.00
Chronic Kidney Disease, Stage 5 (HCC136)	103,556	2%	52.15	-0.45	0.00
Chronic Kidney Disease, Moderate (Stage 3) (HCC138)	407,330	8%	53.62	0.17	0.00
Hip Fracture/Dislocation (HCC170)	6,070	0%	52.99	0.04	0.64
Complications of Specified Implanted Device or Graft (HCC176)	29,235	1%	54.60	0.31	0.00
Amputation Status, Lower Limb/Amputation Complications (HCC189)	34,343	1%	51.99	0.37	0.00
Intercept	.	.	.	34.17	0.00

Exhibit A-2. C-Statistics for Imputation Models across GG Items at SOC/ROC and Discharge, CY 2022

Item/Activity	Description	Assessment Timing	C Statistic
GG0130A	Eating	SOC/ROC	0.85
		Discharge	0.94
GG0130B	Oral Hygiene	SOC/ROC	0.87
		Discharge	0.97
GG0130C	Toileting Hygiene	SOC/ROC	0.92
		Discharge	0.97
GG0170A	Roll left/right	SOC/ROC	0.93
		Discharge	0.98
GG0170C	Lying to sit on Side of bed	SOC/ROC	0.97
		Discharge	0.98
GG0170D	Sit to stand	SOC/ROC	0.97
		Discharge	0.98
GG0170E	Chair to bed trans.	SOC/ROC	0.98
		Discharge	0.99
GG0170F	Toilet transfer	SOC/ROC	0.97
		Discharge	0.98
GG0170I	Walk 10'	SOC/ROC	0.95
		Discharge	0.97
GG0170J	Walk 50' w/ 2 turns	SOC/ROC	0.97
		Discharge	0.98
GG0170R	Wheel 50' w/ 2 turns	SOC/ROC	0.83
		Discharge	0.86