Endorsement & Maintenance Full Measure Submission

Instructions: This form can be used as a worksheet to assist you in developing your **Full Measure Submission (FMS)** for a new or maintenance measure. When you have received the approval for full measure submission, navigate to https://p4qm.org/ and log into your PQM account. Once logged in, click "My Account" to go to your dashboard, then scroll to the bottom of the page and select *Approved for Full Measure Submission* from the "Endorsement Cycle Status" drop–down list and click "Apply" to see your measures ready for FMS. To return to an FMS draft in progress, select *Full Measure Submission Draft* from the drop–down list and click "Apply". Click here for more information on the Endorsement & Maintenance measure submission process.

- You must complete all required fields (denoted by *) to submit the final FMS
- You may save a draft of the FMS form before completing all required fields
- If you would like to make changes to information submitted via the Intent to Submit (ITS), you may edit the original content in the FMS form
- Ensure all attachments are 508 compliant, including labeling all tables and figures with alternative text, as appropriate

Required fields vary depending on whether your measure is an electronic Clinical Quality Measure (eCQM), or an initial (new) measure versus a maintenance measure, or for selected other situations. Conditional fields are indicated in this template with brackets before each field (e.g., *[If the measure is an eCQM]* **Attach MAT Output** *).

Substantive updates to the FMS form for the Fall 2024 cycle:

- Fields moved from the FMS form to the Intent to Submit (ITS) form: Section 6.1, Use (6.1.1 Current Status; 6.1.2–6.1.3 Current/Planned Use; 6.1.4 Program Details)

Section 1. Measure Specifications

[NOTE: Items 1.1–1.9, 1.14, 1.15, 1.15d, and 1.15e were entered in the ITS, and can be edited in the FMS]

1.10 Measure Rationale *

Provide a rationale for why measured entities should report this measure, including how the measure will improve the quality of care for patients and/or any associated health care costs, and what are the benefits or improvements in quality envisioned by use of this measure.

Measuring functional status of home health patients can provide valuable information about a home health agency's (HHAs) quality of care. A patient's functional status may be associated

with adverse health outcomes such as falls, fractures, exacerbation of chronic conditions, and a higher risk of readmissions following home care. Predictors of poorer recovery in function include greater age, complications after hospital discharge, and residence in a nursing home. 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 have shown significantly improved function. Home health services, delivered by a registered nurse positively impacted patient quality of life 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.

The Cross-Setting Discharge Function Score for HH 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 Discharge Function Score for HH 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 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. The Cross-Setting Discharge Function Score for HH measure adds value to the HH Quality Reporting Program (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 activity not attempted codes.

1.11 Measure Webpage *

Provide a URL to a webpage, specific for this measure, containing current detailed specifications, including code lists, risk model details, and supplemental materials. Do not enter a URL to a home page or to general information. The webpage must be publicly accessible. If no URL is available, copy and paste this example: http://example.com.

Home Health Quality Measures | CMS (https://www.cms.gov/medicare/quality/home-health/home-health-quality-measures)

1.12 *[If the measure is an eCQM]* Attach MAT Output

Attach the zipped output from the Measure Authoring Tool (MAT). If you did not use the MAT, please contact PQM Support. Use the measure specification fields (e.g., 1.14a – 1.15c) for the plain–language description of the specifications. One file only; 256 MB limit; Allowed file types: .zip. N/A

1.13 Attach Data Dictionary

Attach a data dictionary, code table, and/or value sets (include variables in the final risk model or stratification plan, if applicable). Attachment should include variables used in the final risk model and/or stratification, if applicable.
One file only; 256 MB limit; Allowed file type: .xls; .xlsx; .csv (please clearly label sheets).

☒ **1.13a Data dictionary not attached**

I attest that all information will be provided in relevant fields where code and/or value sets are needed (e.g., 1.14a – 1.15b).

1.14a Numerator Details *

Provide details needed to calculate the numerator. All information required to identify and calculate the cases from the target population (denominator) with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets. If your list of codes with descriptors is greater than will fit in this text box, you must attach an Excel or csv file in the previous question. If the numerator includes a list (or lists) individual codes with descriptors that exceeds one page, please provide this information in an xls; .xlsx; .csv file as part of the data dictionary attachment.

The numerator is the number of quality episodes during the reporting period in which the observed discharge function score (Section 1) for select standardized functional items/activities is equal to or greater than the expected discharge function score (Section 2).

Section 1. The observed discharge function score is the sum of individual function activities at discharge (see Exhibit 1). The section in each PAC assessment instrument titled Section GG, Functional Ability and Goals, includes standardized patient assessment data elements that measure mobility and self-care functional status. The Cross-Setting Discharge Function Score for HH focuses on these standardized functional activities that are currently available across all PAC settings (Exhibit 1). Valid responses for the standardized functional items/activities are reported in Exhibit 2.

Exhibit 1.Standardized Function Self-Care and Mobility 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

Exhibit 2. Standardized Function Items Response

Category	Standardized Function 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 06 (independent) and 01 (dependent), use the code as the score for that activity.

Step 2: If the code for an activity is 07, 09, 10, 88, dashed (-), skipped (^), or missing, then the score for that activity is estimated with statistical imputation (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.

The definition of patients who are wheelchair users is specified in the Technical Report here: Home Health Discharge Function Technical Report - March 2024 (cms.gov). Different locomotion activities are used if the patient is a wheelchair user 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.

Section 2. 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 4.4 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.

1.15a Denominator Details *

Provide details needed to calculate the denominator. All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets. If the list(s) of individual codes with descriptors exceeds one page, please provide this information in an Excel or .csv file as part of the data dictionary attachment.

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. 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.

1.15b Denominator Exclusions *

Briefly describe exclusions from the denominator cases, if any. Enter "None" if the measure does not have denominator exclusions.

A Medicare Part A episode-level record is excluded if:

1) Patient had an incomplete stay:

- Length of stay is less than 3 days
- Died while in HH (Item M0100 equal to "08")
- Discharge destination indicates the patient had a medical emergency (Item M0100 equal to "06" or "07")

2) Patient has the following medical conditions: Coma, persistent vegetative state, complete tetraplegia, locked-in syndrome, severe anoxic brain damage, cerebral edema or compression of brain (must have a valid diagnosis in Items M1021 and M1023 and Item M1700 equal to "04").

3) Patient is younger than age 18

4) Patient is discharged to hospice

1.15c Denominator Exclusions Details *

Provide details needed to calculate denominator exclusions. Enter "None" if the measure does not have denominator exclusions. All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets. If the list(s) of codes with descriptors exceeds one page, please provide this information in an Excel or .csv file as part of the data dictionary attachment.

In addition to the details given above in response to 1.15b, the following details also inform the application of exclusion criteria.
- Patient is discharged to hospice: item M2420 equal to “03”.
- ICD-10-CM codes (**see attachment in section 1.13**) for coma, complete tetraplegia, locked-in state, persistent vegetative state, severe anoxic brain damage, edema, or compression, severe brain damage

1.16 Type of Score *
Select the most relevant type of score.

☐ Categorical, e.g., yes/no
☐ Continuous variable, e.g., average
☐ Count
☒ Rate/proportion
☐ Composite scale
☐ Other scoring method

1.16a Describe other scoring method *
N/A

1.17 *[If Measure Type (1.5) IS NOT “Cost/Resource Use”]* **Measure Score Interpretation** *
Select the appropriate interpretation of the measure score

☒ Better quality = Higher score
☐ Better quality = Lower score
☐ Better quality = Score within a defined interval
☐ Passing score defines better quality
☐ Other

1.17a Describe Other measure score interpretation *
N/A

1.17 *[If Measure Type (1.5) IS “Cost/Resource Use”]* **Select the type of cost measure** * N/A

☐ Per capita (population– or patient–based)
☐ Per episode
☐ Per procedure
☐ Other

1.17a Specify other cost measure *
N/A

1.18 Calculation of Measure Score *

Diagram or describe the calculation of the measure score as an ordered sequence of steps. Identify the denominator, denominator exclusions (if any), numerator, time period of data collection, risk adjustment and/or stratification, and any other calculations.

The Cross-Setting Discharge Function Score for HH measure 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 HH quality episode begins with either a SOC (start of care) or ROC (resumption of care) and ends with an EOC (end of care) event (a transfer, death, or discharge) for a patient regardless of the length of time between the start and ending events. A higher score indicates better performance in functional outcomes. For each HH quality episode, observed discharge function score and expected discharge function score are determined. For each HHA, the Discharge Function Score is the proportion of quality episodes where the observed discharge function score is greater than or equal to the expected discharge function score.

The Cross-Setting Discharge Function Score for HH focuses on the standardized functional assessment items listed in Exhibit 3 (same as Exhibit 1) that are currently available across all PAC settings.

Exhibit 3. Standardized Function Self-Care and Mobility 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 standardized functional items are reported in Exhibit 4 (same as Exhibit 2).

Exhibit 4. Standardized Function Items Responses

Category	Standardized Function 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

Category	Standardized Function Items Response	Response Description
	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 process for calculating the Cross-Setting Discharge Function Score for HH measure can be divided into two phases. In the first phase, the standardized functional items at SOC/ROC and at discharge that have an Activity Not Attempted (ANA) code of 07, 09, 10, or 88, a dash (-), or a skip (^) recorded (hereafter referred to as NA) are estimated with statistical imputation methods. The estimation models include the predictors used in risk adjustment and covariates for scores on other standardized functional items. Notably, the estimation process uses all standardized functional items available in HH to estimate the NA scores for the subset of standardized functional items used for the Discharge Function Score numerator. See Appendix in Section 7 for more details on the estimation process. In the second phase, the calculation of Discharge Function Score continues. The steps below describe how to calculate the Discharge Function Score.

Step 1: For each HH quality episode, calculate the observed discharge function score by summing the individual standardized functional items. If the standardized functional item has a score of 01 − 06, then use the score for that item. If the standardized functional item has an NA value recorded, then use the imputed score.

A patient is determined to be a wheelchair user if (i) Walk 10 Feet (GG0170I) has an 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.

For patients who are wheelchair users, the observed discharge function score is calculated as sum(GG0130A, GG0130B, GG0130C, GG0170A, GG0170C, GG0170D, GG0170E, GG0170F, (2×GG0170R)). For all other patients, the observed discharge function score is calculated as sum(GG0130A, GG0130B, GG0130C, GG0170A, GG0170C, GG0170D, GG0170E, GG0170F, GG0170I, GG0170J).

Since there are 10 standardized functional items included in the observed discharge function score, each patient's total observed discharge score will range from 10 – 60.

Step 2: Identify excluded HH quality episodes. Excluded HH quality episodes are those that end in a transfer, death at home, or that are less than three days. Also excluded are HH quality episodes where the patient has certain medical conditions, including a primary or other diagnosis indicating coma, persistent vegetative state, complete tetraplegia, locked-in state, severe anoxic brain damage, cerebral edema, or compression of the brain. Finally, HH quality episodes where the patient is discharged to hospice (home or institutional facility) are also excluded.

Step 3: For each HH quality episode, calculate the expected discharge function score. The risk adjustment model is an ordinary least squares linear regression model, which estimates the relationship between discharge function score and a set of risk adjustors.

The risk adjustment model is run on all HHA quality 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 quality episode at SOC/ROC.

$$Expected\ Discharge\ Function\ Score = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n$$

where $x_1 - x_n$ are the risk adjustors.

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.

1.18a Attach measure score calculation diagram

Attach a measure score calculation diagram, if desired.

One file only; 256 MB limit; Allowed file types: .pdf; .jpg; .png.

1.19 Measure Stratification Details *

Provide all information required to stratify the measure results, if necessary. Include the stratification variables, definitions, code/value sets, and if appropriate, the risk–model covariates

and coefficients for the clinically–adjusted version of the measure. If the list(s) of codes with descriptors exceeds one page, please provide this information in an Excel or .csv file as part of the data dictionary attachment. If the measure is not stratified, please state "The measure is not stratified." If the information is included within the data dictionary attachment, please state "See data dictionary attachment."

The measure is not stratified (N/A).

1.20 Testing Data Sources *

Select the data sources for which you have tested and specified the measure. Choose all that apply.

☐ Administrative Data
☐ Claims Data
☐ Electronic Health Records
☐ Paper Patient Medical Records
☐ Registries
☒ Standardized Patient Assessments
☐ Patient–Reported Data and/or Survey Data *[Answer questions 1.21–1.24]*
☐ Non–Medical Data
☐ Other Data Source

1.20a Specify other data source *

N/A

1.21 *[If "Patient–Reported Data and/or Survey Data" was selected above]* Patient reported data collection tools N/A

Choose one (1.21a or 1.21b). If the measure requires patient–reported data to collect stratification and/or risk adjustment variables, please include this information as well.

1.21a Data Source URL(s)

Provide link to the survey, tool, questionnaire, or scale used as a data source for your measure. This must be an external URL such as http://example.com. If no URL is available, copy and paste the example: http://example.com. Click "Add Another Item" to enter multiple URLs.

1.21b Attach Data Collection Tool(s)

Attach the survey, tool, questionnaire, or scale used as a data source for your measure.

One file only; 256 MB limit; Allowed type: .zip.

1.22 *[If "Patient–Reported Data and/or Survey Data" was selected for 1.20]* Proxy Responses * N/A

Are proxy responses allowed?

☐ No ☐ Yes

1.23 *[If "Patient–Reported Data and/or Survey Data" was selected for 1.20]* **Survey Respondent** * N/A

Please indicate the respondent for your survey, tool, questionnaire, or scale. Select all that apply.

☐ Patient
☐ Family or other caregiver
☐ Clinician
☐ Other

1.23a Specify other survey respondent *

1.24 *[If "Patient–Reported Data and/or Survey Data" was selected for 1.20]* **Data Collection and Response Rate** *

For survey/patient–reported data, provide instructions for data collection (e.g., modes of collection, languages of administration), including disclosing minimum response rates and guidance on improving response rates. In addition, specify how to calculate response rates for reporting with performance measure results.

N/A

1.25 Data Sources *

Identify the specific data source(s), other than or in addition to any patient–reported data and/or survey data collection instrument(s) indicated for the measure. For example, provide the name of the database, clinical registry, etc. and describe how the data are collected. Please discuss any data feasibility, reliability, and/or validity challenges and how this has been mitigated.

The data source used is Outcome and Assessment Information Set also known as OASIS. Home Health Quality Measures | CMS (https://www.cms.gov/medicare/quality/home-health/home-health-quality-measures)

1.26 Minimum Sample Size *

Indicate whether the measure has a minimum sample size to calculate the performance score and provide any instructions needed for obtaining the sample and guidance on minimal sample size.

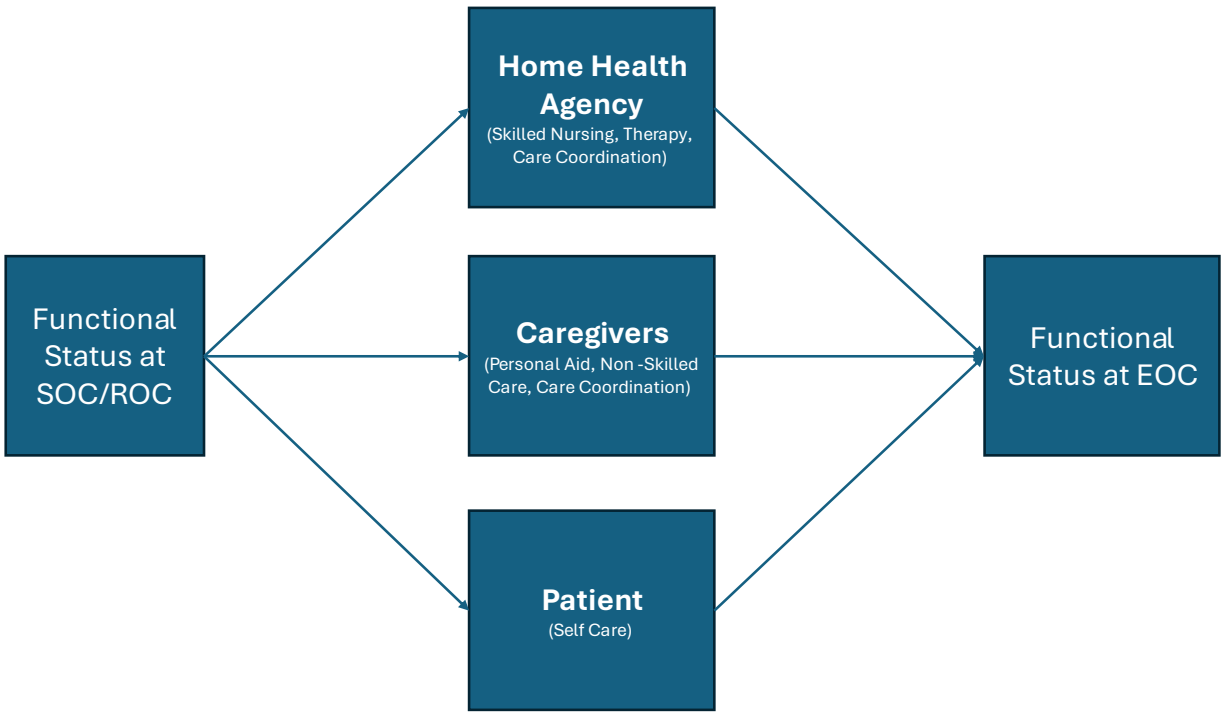
At least 20 eligible quality episodes are required for the Cross-Setting Discharge Function Score for HH measure in the reporting period. In CY 2023, 80.0% (n=8,093) of all HHAs (n=10,122) met this threshold and accounted for 99.7% (n=5,153,932) of all eligible quality episodes among all providers.

Section 2. Importance

2.1 Attach Logic Model *

Attach a logic model depicting the relationship between structures and processes and the desired outcome. Briefly describe the steps between the health care structures and processes (e.g., interventions, or services) and the desired health outcome(s). Identify the relationships among the inputs and resources available to create and deliver an intervention, the activities the intervention offers, and the expected results (i.e., desired outcome). The relationships in the diagram should be easily understood by general, non–technical audiences. Indicate the structure, process, or outcome being measured.

One file only; 256 MB limit; Allowed file types: .pdf; .doc; .docx.



2.2 Evidence of Measure Importance *

Summarize evidence of the measure's importance from the literature, linking the structure/process/intermediate outcome to the desired health outcome. Please provide references for supporting evidence.

Measuring functional status of home health care (HHC) patients can provide valuable information about an HHA's quality of care. Impaired physical functioning is associated with increased healthcare utilization and increased costs to the health care system[1]. It is a well-established risk factor for poor health outcomes including: nursing home admission[2], higher risk of falls and falls-related hip fracture and death,[3,4] greater risk of undernutrition,[5] higher emergency department admissions,[6] higher prevalence of hypertension and diabetes[7], and a higher risk of feelings of loneliness among older adults[8]. Findings from studies specific to the

home healthcare setting have also established that impaired physical function is related to higher risk of infection in HHC patients[9] and a higher risk of potentially preventable hospitalization among HHC beneficiaries[10]. Findings from studies conducted among HHC beneficiaries with dementia indicated that in addition to having the greatest effect on risk for a potentially preventable hospitalization, physical function deficits were also associated with decreased likelihood of successful discharge to community after HHC for those patients[11-13].

Home health care can positively affect functional and other health outcomes. The delivery of home-based physical therapy and skilled nursing provided as part of HHC are associated with improved physical function among HHC beneficiaries[14-16], lower risks of rehospitalization[17], and improved cardiovascular health and blood pressure management[18]. In stroke patients, home based rehabilitation programs administered by home health clinicians significantly improved function.[19] 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.[20] 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[21] 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[22]. Among HHC patients with dementia, physical therapy services increased the likelihood of successful discharge to community[13]. Managing and improving physical function among HHC beneficiaries has also been shown to have a positive effect on caregivers and caregiver support, which is essential to patient well-being and improvement[23].

Better understanding of discharge functional status and the role HHC plays in addressing it can lead to better opportunities to target efforts to improve care for beneficiaries[24]. Current predictors of poorer recovery in function include greater age, complications after hospital discharge, and residence in a nursing home[25]. Measurement of discharge function in HHC can lead to better understanding of additional beneficiary characteristics that may be associated with poorer functional recovery in HHC and how to best estimate the appropriate functional outcome expectations for home health patients based on their personal characteristics and health status. Measurement of discharge function in HHC can facilitate identification of risk factors and better understanding of how home health interventions aimed at improving physical function and discharge performance may be related to decreasing health and safety risks for HHC beneficiaries[10]. Measurement of discharge function can also be used to further explore dose related responses to HHC delivery- such as how specific HHC services (PT, OT, SN) and the number of visits received by a HHC patient may be related to improvements in function and the quality of a home health agency.[16,17,26.] Measurement of discharge function can also be an important determinate for uncovering health care disparities among vulnerable populations in home health care. Racial/ethnic minority status and dementia are associated with less functional improvement in HH. [27,28] Measurement of discharge function and further exploration of related factors can help to clarify the underlying mechanisms that are causing these disparities and how HHC can develop interventions to address them.

The Cross-Setting Discharge Function Score for HH 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 Discharge Function Score for HH 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 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. Physical function performance has been shown to be related to risk for hospitalization at discharge in post-acute care settings and there is a need to determine effective strategies of maintaining and facilitating functional performance across post-acute settings to optimize long-term patient outcomes [29]. The Cross-Setting Discharge Function Score for HH 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. Ultimately, improved mobility and physical function and the prevention of functional decline for HHC beneficiaries has the power to improve their health care status and quality of life far beyond the home health care episode. By maintaining or improving function for HHC beneficiaries, it decreases their risk for hospitalization, improves other health outcomes, and decreases burden on caregivers who will continuing to support their loved one after the end of a HHC episode.

1. Cheng, Y., Goodin, A. J., Pahor, M., Manini, T., & Brown, J. D. (2020). Healthcare Utilization and Physical Functioning in Older Adults in the United States. *Journal of the American Geriatrics Society*, *68*(2), 266–271. https://doi.org/10.1111/jgs.16260
2. 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.
3. 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.
4. 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.
5. [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).
6. 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.
7. 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.
8. Marika Moeyersons, Kristel De Vliegher, Brooke Huyghe, Sacha De Groof, Koen Milisen, & Bernadette Dierckx de Casterlé. (2022). 'Living in a shrinking world'—The experience

of loneliness among community-dwelling older people with reduced mobility: a qualitative grounded theory approach. *BMC Geriatrics*, *22*(1), 1–13. https://doi.org/10.1186/s12877-022-02998-5

9. Shang, J., Wang, J., Adams, V., & Ma, C. (2020). Risk factors for infection in home health care: Analysis of national Outcome and Assessment Information Set data. *Research in Nursing & Health*, *43*(4), 373–386. https://doi.org/10.1002/nur.22053
10. 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.
11. Knox, S., Downer, B., Haas, A., Middleton, A., & Ottenbacher, K. J. (2020). Function and Caregiver Support Associated With Readmissions During Home Health for Individuals With Dementia. *Archives of Physical Medicine & Rehabilitation*, *101*(6), 1009–1016. https://doi.org/10.1016/j.apmr.2019.12.021
12. Knox, S., Downer, B., Haas, A., & Ottenbacher, K. J. (2021). Mobility and Self-Care are Associated With Discharge to Community After Home Health for People With Dementia. *Journal of the American Medical Directors Association*, *22*(7), 1493–1499. https://doi.org/10.1016/j.jamda.2020.12.014
13. Knox, S., Downer, B., Haas, A., & Ottenbacher, K. J. (2022). Home health utilization association with discharge to community for people with dementia. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *8*(1), e12341.
14. Bowles KH, McDonald M, Barrón Y, Kennedy E, O'Connor M, & Mikkelsen M. (2021). Surviving COVID-19 After Hospital Discharge: Symptom, Functional, and Adverse Outcomes of Home Health Recipients. *Annals of Internal Medicine*, *174*(3), 316–325. https://doi.org/10.7326/M20-5206
15. Cook, T., & Mendonca, R. (2017). The Effectiveness of a Comprehensive Home Health Program for Functional Independence of Older Adults in the Home. *American Journal of Occupational Therapy*, *71*, 238. https://doi.org/10.5014/ajot.2017.71S1-PO2018
16. LeDoux, C. V., Lindrooth, R. C., Seidler, K. J., Falvey, J. R., & Stevens-Lapsley, J. E. (2020). The Impact of Home Health Physical Therapy on Medicare Beneficiaries With a Primary Diagnosis of Dementia. *Journal of the American Geriatrics Society*, *68*(4), 867–871. https://doi.org/10.1111/jgs.16307
17. Wang, J., Liebel, D. V., Yu, F., Caprio, T. V., & Shang, J. (2019). Inverse Dose-Response Relationship Between Home Health Care Services and Rehospitalization in Older Adults. *Journal of the American Medical Directors Association*, *20*(6), 736–742. https://doi.org/10.1016/j.jamda.2018.10.021
18. Florindo, M., Gregório, J., & Rodrigues, L. M. (2024). Lower-Limb Perfusion and Cardiovascular Physiology Are Significantly Improved in Non-Healthy Aged Adults by Regular Home-Based Physical Activities—An Exploratory Study. *Life (2075-1729)*, *14*(2), 241. https://doi.org/10.3390/life14020241
19. 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.
20. 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.
21. 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.
22. Marek, K.D., Popejoy, I., Petroski, G. et al. (2005). Clinical outcomes of aging in place. Nurs Res; 54:202–211.

23. Chase, J.-A. D., Russell, D., Rice, M., Abbott, C., Bowles, K. H., & Mehr, D. R. (2019). Caregivers' Perceptions Managing Functional Needs Among Older Adults Receiving Post-Acute Home Health Care. *Research in Gerontological Nursing*, *12*(4), 174–183. https://doi.org/10.3928/19404921-20190319-01
24. Courtright, K. R., Jordan, L., Murtaugh, C. M., Barrón, Y., Deb, P., Moore, S., Bowles, K. H., & Mikkelsen, M. E. (2020). Risk Factors for Long-term Mortality and Patterns of End-of-Life Care Among Medicare Sepsis Survivors Discharged to Home Health Care. *JAMA Network Open*, *3*(2), e200038. https://doi.org/10.1001/jamanetworkopen.2020.0038
25. 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.
26. Falvey, J. R., Bade, M. J., Forster, J. E., Burke, R. E., Jennings, J. M., Nuccio, E., & Stevens-Lapsley, J. E. (2018). Home-Health-Care Physical Therapy Improves Early Functional Recovery of Medicare Beneficiaries After Total Knee Arthroplasty. *The Journal of Bone and Joint Surgery. American Volume*, *100*(20), 1728–1734. https://doi.org/10.2106/JBJS.17.01667
27. Chase, J.-A. D., Huang, L., Russell, D., Hanlon, A., O, C. M., Robinson, K. M., & Bowles, K. H. (2018). Racial/ethnic disparities in disability outcomes among post-acute home care patients. *Journal of Aging & Health*, *30*(9), 1406–1426. https://doi.org/10.1177/0898264317717851
28. Wang, J., Yu, F., Cai, X., Caprio, T. V., & Li, Y. (2020). Functional outcome in home health: Do racial and ethnic minority patients with dementia fare worse? *PLoS ONE*, *15*(5), 1–17. https://doi.org/10.1371/journal.pone.0233650
29. Li, C.-Y., Haas, A., Pritchard, K. T., Karmarkar, A., Kuo, Y.-F., Hreha, K., & Ottenbacher, K. J. (2021). Functional Status Across Post-Acute Settings is Associated With 30-Day and 90-Day Hospital Readmissions. *Journal of the American Medical Directors Association*, *22*(12), 2447–2453. https://doi.org/10.1016/j.jamda.2021.07.039

2.3 *[If initial endorsement]* **Anticipated Impact** *

If implemented, what is the measure's anticipated impact on the desired outcomes, such as those listed in the logic model? Please cite evidence to identify adverse events and costs avoided and provide references. Describe how the benefits of the measure's impact will outweigh any potential unintended consequences.

Home health care (HHC) can positively affect functional and other health outcomes. The delivery of home-based physical therapy and skilled nursing provided as part of HH are associated with improved physical function among HH beneficiaries[14-16], lower risks of rehospitalization[17], and improved cardiovascular health and blood pressure management[18]. Thus, the Cross-Setting Discharge Function Score for HH measure can improve patient outcomes in post-acute care (PAC) by promoting functional independence, reducing adverse events, and lowering healthcare costs. Managing and improving physical function among HHC beneficiaries has also been shown to have a positive effect on caregivers and caregiver support, which is essential to patient well-being and improvement[23].

The Cross-Setting Discharge Function Score for HH measure determines how successful each Home Health Agency (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 Discharge Function Score for HH 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 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. In stroke patients, home based rehabilitation programs administered by home health clinicians significantly improved function.[19] 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.[20] 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[21] 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[22].

The Cross-Setting Discharge Function Score for HH 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.

One concern about unintended consequences with the Cross-Setting Discharge Function Score for HH is that the measure may lead HHAs to selectively enroll patients, either by encouraging or avoiding admission of certain types of patients and patients with certain characteristics. To address this, providers' performance is evaluated among their peers after adjusting for difference in patient case-mix across HHAs. The risk adjustment methodology applied to this measure will help mitigate providers' incentive to selectively enroll patients. The variables included in the risk adjustment model are designed to capture patient characteristics that are associated with discharge functional status. Therefore, providers' performance on this measure will be adjusted for the characteristics of their patient population and "level the playing field" across providers. The detailed risk-adjustment strategy will be publicly available, allowing providers to understand that those who provide care for more "high risk" patients are not at a disadvantage given their patient case-mix. See the Technical Report for more details on the risk adjustment methodology.

Another potential concern about the Cross-Setting Discharge Function Score for HH measure could be that it focuses on a subset of the available standardized functional self-care and mobility items in HH. If the items are not included in this publicly reported measure, it could reduce the incentive to complete those items and could result in higher levels of ANAs. However, the standardized functional items excluded from the Cross-Setting Discharge Function Score for HH measure may be used in the future for the HH prospective payment system to calculate payment for HHAs and are included in the statistical imputation models for the Cross-Setting Discharge Function Score for HH measure. Together, these circumstances should provide an incentive for continued reporting of these standardized functional items.

Another possibility related to increased ANA rates is that providers could strategically code ANAs in an attempt to game the estimated values from the statistical imputation models. For instance, HHAs could record ANA codes for patients who did not improve by discharge if the discharge estimation models would predict higher scores based on that patient's characteristics. However, this type of gaming, where providers are determining in real-time which patients would perform better with statistical estimation than a true discharge score, would require sophisticated understanding and application of the estimation methodology.

The Cross-Setting Discharge Function Score for HH measure will be monitored to identify unintended consequences, including patient selection patterns or changes in ANA coding, which could lead to future re-specification of the measure as needed.

13. Knox, S., Downer, B., Haas, A., & Ottenbacher, K. J. (2022). Home health utilization association with discharge to community for people with dementia. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *8*(1), e12341.
14. Bowles KH, McDonald M, Barrón Y, Kennedy E, O'Connor M, & Mikkelsen M. (2021). Surviving COVID-19 After Hospital Discharge: Symptom, Functional, and Adverse Outcomes of Home Health Recipients. *Annals of Internal Medicine*, *174*(3), 316–325. https://doi.org/10.7326/M20-5206
15. Cook, T., & Mendonca, R. (2017). The Effectiveness of a Comprehensive Home Health Program for Functional Independence of Older Adults in the Home. *American Journal of Occupational Therapy*, *71*, 238. https://doi.org/10.5014/ajot.2017.71S1-PO2018
16. LeDoux, C. V., Lindrooth, R. C., Seidler, K. J., Falvey, J. R., & Stevens-Lapsley, J. E. (2020). The Impact of Home Health Physical Therapy on Medicare Beneficiaries With a Primary Diagnosis of Dementia. *Journal of the American Geriatrics Society*, *68*(4), 867–871. https://doi.org/10.1111/jgs.16307
17. Wang, J., Liebel, D. V., Yu, F., Caprio, T. V., & Shang, J. (2019). Inverse Dose-Response Relationship Between Home Health Care Services and Rehospitalization in Older Adults. *Journal of the American Medical Directors Association*, *20*(6), 736–742. https://doi.org/10.1016/j.jamda.2018.10.021
18. Florindo, M., Gregório, J., & Rodrigues, L. M. (2024). Lower-Limb Perfusion and Cardiovascular Physiology Are Significantly Improved in Non-Healthy Aged Adults by Regular Home-Based Physical Activities—An Exploratory Study. *Life (2075-1729)*, *14*(2), 241. https://doi.org/10.3390/life14020241
19. 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.
20. 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.
21. 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.
22. Marek, K.D., Popejoy, I., Petroski, G. et al. (2005). Clinical outcomes of aging in place. Nurs Res; 54:202–211.
23. Chase, J.-A. D., Russell, D., Rice, M., Abbott, C., Bowles, K. H., & Mehr, D. R. (2019). Caregivers' Perceptions Managing Functional Needs Among Older Adults Receiving Post-Acute Home Health Care. *Research in Gerontological Nursing*, *12*(4), 174–183. https://doi.org/10.3928/19404921-20190319-01

2.4 Performance Gap

If available, provide evidence of performance gap or measurement gap by providing performance scores on the measure as specified at the specified level(s) of analysis. Please include mean, minimum, maximum, and scores by deciles by using the table below or upload an attachment. In the text field here, describe the data source, including number of measured entities, number of patients, dates of data. If a sample was used, provide characteristics of the entities included. If performance scores are unavailable for the measure, please explain.

There is evidence of a performance gap and variability in performance for this quality measure. Table 1 below reports on data for all 8,093 HHAs that met the minimum threshold of quality episodes for public reporting of the Cross-Setting Discharge Function Score for HH measure (≥20) in the twelve-month reporting period of CY 2023. The data reported in Table 1 provides evidence of a performance gap among providers as performance ranges from the minimum possible score of 0 for 50 HHAs to the maximum possible score of 100 achieved by 20 HHAs, and a mean performance score of 57.4 among all reportable providers.

Table 1 Performance Scores by Decile

Enter overall mean, minimum, and maximum performance scores, along with the count of measured entities and persons/encounters/episodes. Organize entities into deciles by performance scores from 1 (low scores) to 10 (high scores), noting that "high" refers to magnitude, not quality. Provide mean performance scores, number of entities (total) and number of persons/encounters/episodes (total) for entities assigned to each decile.

	Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max
Mean Performance Score	57.4	0.0	14.8	36.4	48.0	54.9	59.5	63.3	66.6	70.2	74.9	84.7	100.0
N of Entities	8,093	50	807	808	812	809	805	812	812	809	809	810	20
N of Persons / Encounters / Episodes	5,153,932	2,886	108,150	235,962	309,313	483,183	813,523	765,036	834,311	683,068	632,220	289,166	1,153

2.4a Attach Performance Gap Results

If needed, you may attach additional performance gap results here. If submitting an attachment rather than entering results in Table 1 above, please enter the overall mean, minimum, and maximum scores, and mean scores by decile. Enter the number of measured entities and persons/encounters/episodes overall and within each decile. Please ensure all attachments are 508 compliant, all tables and figures are labeled with alternative text, as appropriate. Please clearly refer to any results within your attachment within the relevant text fields of this measure submission form. One file only; 256 MB limit; Allowed types: .zip, .pdf, .docx, .xls, .xlsx

2.5 *[If initial endorsement]* Health Care Quality Landscape *

Please explain why existing measures/quality improvement programs are insufficient for addressing this health care need.

The Cross-Setting Discharge Function Score for HH measure was developed based on input obtained during two Technical Expert Panel (TEP) meetings (July 2021 and January 2022). This measure is designed to address the functional status quality measure domain for the home

health as outlined by the Improving Medicare Post-Acute Care Transformation Act of 2014 (the IMPACT Act).[1] During these meetings, panelists expressed that:

1. The HH QRP would benefit from having a cross-setting functional outcome measure to use instead function process measure (Application of Percent of Long-Term Care Hospital (LTCH) Patients With an Admission and Discharge Functional Assessment and a Care Plan That Addresses Function that was recently removed from the HH QRP. The function process measure was removed because it was topped out, lacking the ability to differentiate provider quality related to patient functional status. The Cross-Setting Discharge Function Score for HH measure has higher variation in provider performance and offers more informative comparisons between HHAs for patients, caregivers, and stakeholders.
2. The Cross-Setting Discharge Function Score for HH measure benefits from being specified to align across PAC settings (IRF, LTCH, SNF, HHA). Panelists reviewed comparisons between provider scores and model fit and found that the narrower set of standardized functional items provides similar capture of functional status.[2]
3. The Activity Not Attempted (ANA) codes are used frequently on assessments for certain standardized functional items, and statistical imputation should be used as the method to estimate resulting missing item scores.

The Cross-Setting Discharge Function Score for HH measure will be implemented in the HH QRP and HH VBP Program in CY 2025. There doesn't currently exist a cross setting post-acute care measure of discharge function.

[1] https://www.govtrack.us/congress/bills/113/hr4994

[2] https://www.cms.gov/sites/default/files/2022-04/PAC-Function-TEP-Summary-Report-Jul2021.pdf

2.6 Meaningfulness to Target Population *

Provide evidence the target population (e.g., patients) values the measured outcome, process, or structure, and finds it meaningful. Please describe how and from whom you obtained input.

Functional status, including ability to perform daily activities, is important from patient and caregiver perspectives, with functional goal-setting being an important focus of patient- and family- centered care. For the majority of patients in post-acute care, promoting functional independence and setting functional goals to facilitate return to community living is a primary goal of care. For patients receiving home health services, functional assessment and goal-setting are also a primary focus to attain independent functioning in the home and community, return to or surpass prior level of functioning, maintain current level of functioning, or slow the process of functional decline. In HH settings, promoting physical function is important to mitigate functional deterioration, morbidity, and potential medical complications from disease processes and hospitalization. From a caregiver perspective, focus on functional status and functional goal-setting is important to reduce caregiver burden, and minimize need for assistance at home.

CMS convened a Patient and Family Engagement Listening Session to discuss this measure with patients and their caregivers. 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 facilities made improvements in self-care and mobility.

The Cross-Setting Discharge Function Score for HH directly reflects the priorities of PAC patients, who value functional independence, quality of life, and avoiding rehospitalization or institutionalization. The measure addresses outcomes that patients themselves find meaningful, providing a clear rationale for its adoption and endorsement.

Below are key points supported by peer-reviewed literature:

1. Patients Value Functional Independence:
Studies show that post-acute care patients prioritize functional recovery (e.g., mobility, self-care) as the most important outcome following discharge. Functional independence enables patients to return home and manage daily life without relying on long-term institutional care or home health services.

Source: Graham, J. E., et al. (2016). "Patients' perspectives on discharge from post-acute care settings: Priorities for functional recovery." Archives of Physical Medicine and Rehabilitation.

2. Improved Quality of Life:
Health-related quality of life encompasses patients' physical health perceptions and functional status. Patients who regain independence in activities of daily living report higher satisfaction with their health and overall life post-discharge. They value avoiding dependency on caregivers, especially for basic tasks like toileting and dressing.

Source: Greenfield, S., and Nelson, E. (2020). "The influence of functional independence on quality of life in post-acute care patients." Quality of Life Research.

Additionally, researchers exploring patient and consumer perspectives on function have reported that functional status and functional outcomes are important from the patient and consumer perspective (Stineman 2009, Kramer 1997, Kurz 2008). These studies show that patients place a value on their functional outcomes and rehabilitation goals mostly through research that examines how patients can categorize their functional goals in hierarchies of what they perceive as the most important to least important functional outcomes for the purpose of their own quality of life. Stineman's research shows patients and consumers value their functional outcomes although inpatient rehabilitation patients may have different perspectives on what is important for them to gain from their rehabilitation compared to community dwelling consumers. One study, specifically focused on patients undergoing rehabilitation in IRFs (n=79) found that eating was the most valued functional activity for them, followed by bathing, toileting, and bowel/bladder function (Stineman 2009).

3. Reduction in Hospital Readmissions:
Patients view avoiding rehospitalization as crucial to their recovery. Research demonstrates that patients who regain functional independence are less likely to be readmitted, an outcome patients find meaningful as it reduces the emotional and physical stress of hospitalization.

Source: Ouslander, J.G., and Berenson, R.A. (2011). "Reducing Unnecessary Hospitalizations of Nursing Home Residents." The New England Journal of Medicine.

4. Desire to Return Home:
A primary goal for many PAC patients is to return home after their rehabilitation. Being discharged with higher functional ability is highly valued because it enables patients to live in their communities, reducing the need for institutional care or home health services.

Source: Harrison, S., et al. (2017). "Patient priorities in post-acute care: Returning home with functional independence." Journal of Aging & Health.

5. Patients Want to Avoid Institutional Long-Term Care:
Patients fear the loss of autonomy associated with long-term care facilities and express a strong preference for achieving the functional status that allows them to avoid this outcome. Functional independence is a top priority for maintaining control over their living situation.

Source: Kane, R.A. (2001). "Long-Term Care and Patient Preferences: Achieving Independence and Control." The Gerontologist.

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.

Section 3. Feasibility

3.1 Feasibility Assessment *

Describe the feasibility assessment conducted showing you considered the people, tools, tasks, and technologies necessary to implement this measure. For maintenance measures, describe whether feasibility issues due to implementation might have arisen and the near–term (i.e., within one year) mitigation approaches

The feasibility assessment should address:

- *Whether all required data elements are routinely generated and used during care delivery*
- *The extent of any missing data, measure susceptibility to inaccuracies, and the ability to audit data to detect problems*
- *Estimates of the costs or burden of data collection, data entry, and analysis including the impact on clinician workflow, diagnostic thought processes, and patient–physician interaction*
- *Barriers encountered or that could be encountered in implementing the measure specifications, data abstraction, measure calculation, or performance reporting*
- *Ability to collect information without violation of patient confidentiality, including circumstances where measures based on patient surveys or the small number of patients may compromise confidentiality*
- *Identification of unintended consequences*

OASIS data collection and submission is a requirement of the Medicare Home Health Conditions of Participation. Functional assessment is conducted as part of usual clinical practice, and information on functional status used to calculate this measure is recorded in the relevant OASIS items embedded in the agency's clinical assessment. OASIS data are collected by the home health agency during the care episode and submitted electronically to CMS via the Internet Quality Improvement and Evaluation System (iQIES). No issues regarding availability of data, missing data, timing or frequency of data collection, patient confidentiality or implementation have become apparent since OASIS-E was implemented 1/1/2023.

3.2 *[If an eCQM]* Attach Feasibility Scorecard * N/A

Attach your completed feasibility scorecard; please create the scorecard using the approved template [link].

One file only; 256 MB limit; Allowed types: xlsx.

3.3 Feasibility Informed Final Measure *

Describe how the feasibility assessment informed the final measure specifications, indicating any decisions made to adjust the measure in response to feasibility assessment.

OASIS data collection and submission is a requirement of the Medicare Home Health Conditions of Participation.

3.4 Proprietary Information *

Indicate whether your measure or any of its components are proprietary, with or without fees (choose one).

☐ Proprietary measure or components (e.g., risk model, codes), without fees
☐ Proprietary measure or components with fees
☒ Not a proprietary measure and no proprietary components

3.4a *[If any proprietary components for 3.4]* **Fees, Licensing, or Other Requirements** *

Describe any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm).

None

Section 4. Scientific Acceptability

4.1 Data and Samples

4.1.1 Data Used for Testing *

Describe the data used for testing (include dates, sources).

The analyses presented in this form are calculated from data derived from several sources. The primary source of data for the measure is OASIS assessment data from Calendar Year (CY) 2023. The OASIS assessments are combined into HH quality episodes.

A HH quality episode begins with either a SOC (start of care) or ROC (resumption of care) and ends with an EOC (end of care) event (a transfer, death, or discharge) for a patient regardless of the length of time between the start and ending events.

Quality episodes are created by matching SOC/ROC assessments with EOC assessments for a given patient who receives care by a home health agency. A matching pair of assessments is then turned into one quality episode that provides information about the patient collected via the OASIS instrument at the two time points and thus allows for analysis of changes in a patient's health status between the two time points.

All analyses, unless otherwise indicated, were calculated using HH quality episodes from HHAs that met the reportability threshold of at least 20 quality episodes after applying denominator exclusion criteria.

For analyses related to health equity, we also used Medicare administrative data to determine dual eligibility status for Medicare and Medicaid and Area Deprivation Index (ADI) data, derived from American Community Survey data. The ADI is presented as an index ranging from zero to 100, designed to represent neighborhood socioeconomic disadvantage, with 100 representing the most disadvantaged neighborhoods nationwide.

4.1.2 Differences in Data *

If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), clearly identify which data source/sample is used for each aspect of testing, including the years of data used in each. If there are no differences to report, enter "None."

The sample remained the same for all aspects of testing. For testing of differences in performance scores across socio-contextual variables, including race, ethnicity and socio-economics status (see Section 5. Equity), we used additional data sources to incorporate ADI, derived from census data, and dual eligibility for Medicare and Medicaid from CMS administrative data.

4.1.3 Characteristics of Measured Entities *

Describe characteristics of measured entities included in the analysis (e.g., number, size, location, type). If you used a sample, describe how you selected measured entities for inclusion in the sample and the representativeness of the sample.

All testing used HH quality episodes in CY 2023. We had a total of 7,127,963 quality episodes with an episode end date in CY2023 before exclusions. We identified providers that met the reportability threshold of at least twenty quality episodes after applying denominator exclusion criteria. After applying denominator exclusion criteria and the reportability threshold of 20 quality episodes, this testing ultimately included 5,153,932 quality episodes in 8,093 HHAs.

The included HHAs were geographically diverse, with the West – Pacific Census Division containing the largest percentage of HHAs at 23%. The majority of the HHAs were for-profit entities (59%) and located in urban areas (75%). Agency size is presented based on the number of quality episodes. Exactly half of HHAs were medium-sized with 85 to 664 quality episodes, while the other half were evenly divided between small (20 to 84 quality episode) and large (665 to 37,866 quality episodes) in size. Note that providers with less than 20 stays during the 12-month testing period are excluded from analyses presented below.

Exhibit 5 identifies characteristics, including number, percent of episodes and percent of providers, of the publicly reportable home health agencies. Characteristics include size, profit status, rural/urban location, and Census Region/Division.

Exhibit 5. CY 2023 Provider-Level Characteristics among Home Health Agencies Exceeding the Public Reporting Threshold (n>=20)

Characteristic	Category	Publicly Reportable Providers			
		N Episodes	% Episodes	N Providers	% Providers
Overall		**5,153,932**	**100**	**8,093**	**100**
Stay Count	Large (665 - 37866)	1,291,498	25	189	2
	Medium (85 - 664)	2,584,988	50	1,741	22
	Small (20 - 84)	1,277,446	25	6,163	76
Profit Status	For-Profit	2,617,659	51	4,789	59
	Not For-Profit	1,632,005	32	1,317	16
	Government	80,891	2	211	3
	Unknown	837,972	16	1,788	22
Rurality	Rural	681,087	13	1,315	16
	Urban	3,885,590	75	6,195	77
	Unknown	587,255	11	583	7
Region	Northeast - New England	340,803	7	284	4
	Northeast - Middle Atlantic	674,099	13	349	4
	Midwest - East North Central	745,562	15	1,212	15

	Midwest - West North Central	257,269	5	502	6
	South - South Atlantic	1,189,068	23	1,390	17
	South - East South Central	360,813	7	353	4
	South - West South Central	550,846	11	1,472	18
	West - Mountain	248,715	5	581	7
	West - Pacific	726,541	14	1,847	23

4.1.4 Characteristics of Units of the Eligible Population *

Describe characteristics of the patients, encounters, episodes, etc., including numbers and percentages by factors such as age, sex, race, or diagnosis. Provide descriptive statistics separately by each specified level of analysis and data source. If you used a sample, describe how you selected the patients for inclusion in the sample and the representativeness of the sample. If there is a minimum case count used for testing, you must reflect that minimum in the specifications in Minimum Sample Size in Section 1.

All testing used HH quality episodes completed in CY2023. HHA submitted a total of 7,127,963 quality episodes that ended in CY2023. After applying denominator exclusion criteria and applying the reportability threshold of 20 quality episodes, the final sample included 5,153,932 quality episodes in the measure population and testing. For included HH quality episodes, 86.6% were for patients who were over the age of 65 and the majority were female (61.2%) and white (63.1%). Roughly 50% of included quality episodes had an Area Deprivation Index (Neighborhood Atlas - Home (wisc.edu)) of 0–50. The risks of hospitalization among HH patients were varied – 94.7% were taking five or more medications, 67.7% reported exhaustion, 42.4% had a history of falls, and 54.9% had other risks.

Exhibit 6 identifies the patient characteristics of quality episodes treated by publicly reporting home health agencies. Characteristics are reported by race, sex, age, payer source, ADI, health-related social need, and signs or symptoms of risk of hospitalization.

Exhibit 6. CY 2022 Patient Characteristics for Quality Episodes of Care for Home Health Agencies Exceeding the Public Reporting Threshold (n>=20)

Characteristic	Category	PR Providers Eligible Episodes	
		N	%
Overall		5,153,932	100
Race	White	3,253,144	63.1
	Black	492,083	9.5
	Hispanic	385,653	7.5
	Asian	123,161	2.4
	American Indian/Alaska Native	16,254	0.3
	Native Hawaiian/Pacific Islander	9,231	0.2
	Multiple Race/Ethnicity	204,025	4
	No Information Available	1,276,873	24.8
Sex	Male	2,000,573	38.8

	Female	3,153,359	61.2
Age	≤ 54	264,556	5.1
	55-64	429,018	8.3
	65-74	1,485,868	28.8
	75-84	1,744,810	33.9
	85 - 90	782,577	15.2
	> 90	447,103	8.7
Payer	Medicare	4,889,097	94.9
	Dual	134,497	2.6
	Medicaid	399,332	7.7
	Neither Medicare nor Medicaid	0	0
	Unknown	0	0
ADI	>= 85	506,002	13.3
	75-84	360,177	9.5
	50-74	961,058	25.3
	25-49	1,009,640	26.6
	0-24	962,298	25.3
	Missing	1,354,757	26.3
Health Related Social Need/SDOH items	Interpreter need	153,622	3
	Transportation need	384,230	7.5
	Health Literacy need	848,877	16.5
	Social Isolation	117,630	2.3
Risk of hospitalization - signs or symptoms	History of falls	2,183,615	42.4
	Multiple hospitalizations (> 1) in past 6 months	1,655,654	32.1
	Multiple emergency department visits in past 6 months	1,687,576	32.7
	Decline in mental, emotional, or behavioral status, past 3 months	2,292,845	44.5
	History of difficulty complying w/ medical instructions, past 3 months	2,762,008	53.6
	Currently taking 5 or more medications	4,882,935	94.7
	Unintentional weight loss of 10 pounds or more in past 12 months	562,325	10.9
	Currently reports exhaustion	3,488,011	67.7
	Other risks	2,831,100	54.9
	None of the above risks	25,056	0.5

4.2 Reliability

4.2.1 Level(s) of Reliability Testing Conducted *

Choose all that apply.

☒ Person– or encounter–level (i.e., data element) (e.g., inter–abstractor reliability)
☒ Accountable entity–level (e.g., signal–to–noise analysis)
☐ Not applicable/reliability testing not conducted

4.2.1a *Please explain why reliability testing was not conducted*

N/A

4.2.2 *[If reliability testing was conducted]* Method(s) of Reliability Testing *

For each level of reliability testing conducted, describe the method(s) of reliability testing and explain what each tests. Describe the steps, do not just name a method. What type of error does it test? Provide the type of statistical analysis used. Describe proportion of missing data, how missing data was analyzed and/or excluded, and any sensitivity analysis conducted.

Note: Testing at the person or encounter level requires that all critical data elements be tested (not just agreement of one final overall computation for all persons/encounters/episodes). At a minimum, the numerator, denominator, and exclusions must be assessed and reported separately. Prior evidence of reliability of data elements for the data type specified in the measure (e.g., hospital claims) can be used as evidence for those data elements. Prior evidence could include published or unpublished testing that: includes the same data elements, uses the same data type (e.g., claims, chart abstraction), and is conducted on a sample as described above (i.e., representative, adequate numbers, and randomly selected, if possible).

We report testing results throughout this document at two levels: 1) data elements (i.e., items) and the function scale (i.e., summed value derived from item codes) and 2) the computed quality measure result.

To assist the reader in understanding the testing analysis and results, we begin by providing a brief overview of these components of the performance measure:

1. Data Elements:
    a. Clinicians code 11 motor function data elements included in Section GG of each assessment instrument. One is a wheelchair data element used for patients who do not walk as part of the recoding approach. Depending on the context, we sometimes refer to these data elements as “items” or “activities.”
    b. The motor function data elements are collected at the time of admission and discharge using a 6-level rating scale (01 to 06), or activity not attempted codes if, for example, the activity was not attempted due to medical or safety concerns.
    c. Higher scores indicate higher ability (i.e., more independence).
    d. For the performance measure calculation, data element activity not attempted codes and missing data are recoded using statistical imputation to estimate the item score.
    e. A discharge function scale score is created by summing the data element scores, after re-coding. The range of the discharge function score is 10 to 60 units.

   f. For the Cross-Setting Discharge Function Score for HH measure, a score of 10 indicates the patient is dependent on a helper to perform all activities (i.e., data elements) and a score of 60 means the patient is independent on all activities.
2. Calculated Performance Measure Score: The Percentage of HH quality episodes that Meet or Exceed an Expected Discharge Function Score
   a. The calculated performance measure score is the percentage of HH quality episodes that meet or exceed an expected discharge function score within an HHA.
   b. This performance measure estimates the percentage of HH quality episodes that meet or exceed an expected discharge function score.

We use three methods for reliability testing: Cronbach's alpha coefficient, split-sample reliability testing, and signal-to-noise ratio testing.

Cronbach's alpha coefficient assesses the internal consistency of the function scale/instrument scores for each assessment. Internal consistency provides a general assessment of how well the function data elements interrelate within the function scale/instrument. This internal consistency analysis is an indicator of the reliability of the function scale/instrument.

Internal consistency was assessed using the Cronbach's alpha coefficient, which is the average correlation of all possible half-scale divisions. Cronbach's alpha is a statistic frequently calculated when testing instrument or scale psychometrics. The Cronbach's alpha reliability estimate ranges from zero to one, with an estimate of zero indicating that there is no consistency of measurement among the items, and one indicating perfect consistency. Many cutoff criteria exist to determine whether or not a scale shows good consistency or whether the items "hang together" well. General consensus is that Cronbach's alpha should be at least 0.70 for an adequate scale for group-level decisions, and alphas closer to 1 indicate a good scale[1].

Split-sample reliability testing examines the agreement between two performance measure scores for an HHA based on randomly split, independent subsets of patient quality episodes within the same measurement period. We randomly divided each HHA's CY2023 patient quality episodes into halves and calculated performance measure scores for each split-half sample using the same measure specification. We then calculated Shrout-Fleiss[2] intraclass correlation coefficients (ICC[2,1]) between the split-half scores to measure reliability.

Signal-to-noise reliability testing examines the overall reliability of the measure scores by comparing the true effect (the signal) to the error (the noise). We estimated the signal-to-noise ratio in two ways. We first followed the RAND methodology[3] which is reported below in 4.2.3. Then, as a robustness check, we also estimated the ratio by using the sample variance to estimate the provider-to-provider variance.

We performed reliability testing on all HHAs with 20 or more patient quality episodes. These patient quality episodes had complete data.

[1] Aron A, Aron EN *Statistics for Psychology*. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999

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

[3] Adams, J. L. (2009). The reliability of provider profiling: a tutorial.

4.2.3 *[If reliability testing was conducted]* Reliability Testing Results *

Provide the statistical results from reliability testing for each level and type of reliability testing conducted. Where applicable, include results from accountable entity–level reliability testing (e.g., signal–to–noise testing) in the table below.

As reported in the table below, the results for Cronbach's alpha indicate very good inter-rater reliability.

Exhibit 7. Cronbach's Alpha Scale Analysis – Internal Consistency, CY 2023 Data

Item ID	Item Name	Mean	SD	N	Standardized Alpha
GG0130A	Eating	4.83	0.99	3,833,941	0.95
GG0130B	Oral hygiene	4.28	1.05	3,833,941	0.95
GG0130C	Toileting hygiene	3.56	1.11	3,833,941	0.95
GG0170A	Roll left and right	4.14	1.21	3,833,941	0.95
GG0170C	Lying to sitting on side of bed	3.82	1.15	3,833,941	0.94
GG0170D	Sit to stand	3.57	0.98	3,833,941	0.94
GG0170E	Bed-to-chair transfer	3.51	0.97	3,833,941	0.94
GG0170F	Toilet transfer	3.51	1.02	3,833,941	0.94
GG0170I	Walk 10 feet	3.62	0.97	3,833,941	0.94
GG0170J	Walk 50 feet with two turns	3.44	0.99	3,833,941	0.94

As reported in the table below, the results for split-half ICC indicate excellent agreement between performance scores.

Exhibit 8. Reliability Testing Results for Cross-Setting Discharge Function Score for HH, CY 2023 Data

Provider Sample	Stay Count	Provider Count	Split-Half Reliability - ICC	Raw ICC(2,1)	Stays (Min)	Stays (Max)
Overall	**5,153,932**	**8,093**	**0.95**	**0.91**	**20**	**37,866**
Large (665–37,866)	3,938,894	2,022	0.98	0.96	665	37,866
Medium (85–664)	1,117,807	4,037	0.97	0.94	85	664
Small (20–84)	97,231	2,034	0.92	0.86	20	84

As reported in **Table 2** below, the results for signal-to-noise reliability indicate a high proportion of variability in the measure performance scores can be explained by real differences in performance.

4.2.3a *[If reliability testing was conducted]* Attach Additional Reliability Testing Results

If needed, you may attach additional reliability testing results here. Please ensure all attachments are 508 compliant, all tables and figures are labeled with alternative text, as appropriate. Please clearly refer to any results within your attachment within the relevant text fields of this measure submission form.

One file only; 256 MB limit; Allowed types: .zip, .pdf, .docx, .xls, .xlsx

Table 2 *[If accountable entity–level testing was conducted, i.e., if 4.2.1 includes "Accountable Entity–Level")]* **Accountable Entity–Level Reliability Testing Results by Denominator-Target Population Size**

Enter overall mean, minimum, and maximum performance scores, along with the count of measured entities and persons/encounters/episodes. Organize entities into deciles by the entity number of persons/encounters/episodes (denominator) from 1 (smallest N) to 10 (largest N). Provide mean reliability, performance score, number of entities (total) and number of persons/encounters/episodes (total) for entities assigned to each decile. For minimum reliability,

provide reliability value for the entity with the smallest N. For maximum reliability, provide the reliability value for the entity with the largest N.

	Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max
Reliability	0.96	1.00	0.96	0.94	0.94	0.96	0.97	0.97	0.97	0.97	0.97	0.97	1.00
Mean Performance Score	57	0.00	14.82	36.47	48.06	54.85	59.54	63.27	66.64	70.16	74.94	84.72	100.00
N of Entities	8,093	50	807	812	809	808	805	812	812	809	811	808	20
N of Persons / Encounters / Episodes	5,153,932	2,886	108,150	236,215	309,258	482,985	813,523	765,036	834,311	683,068	638,549	282,837	1,153

4.2.4 *[If reliability testing was conducted]* Interpretation of Reliability Results *

Provide your interpretation of the results in terms of demonstrating reliability for each level and type of reliability testing conducted. How do the results support an inference of reliability for the measure?

Split-half Reliability (unit of analysis of providers)
Split-half analysis results indicated strong, positive correlations between the HHAs' randomly divided groups' Computed Quality Measure scores, providing strong evidence of measure reliability with an ICC of 0.95 overall. As shown above in 4.2.3, ICCs were exceptionally strong for providers with higher volume, with ICC of 0.98 among the largest providers (665–37,866 quality episodes)[4].

Signal to Noise (unit of analysis of providers)
Signal to Noise Testing suggests strong reliability across providers, with a reliability statistic ρ of 0.96. Robustness checks in which we calculated the Signal-to-Noise Reliability (VAR) using the sample variance gave an overall statistic of 0.91. Both of these pass the threshold of acceptable reliability.

Cronbach's Alpha (unit of analysis of episodes)
Cronbach's alpha was 0.95 at admission and 0.97 at discharge for patients who do not use wheelchairs and was 0.94 at admission and 0.96 for patients who use wheelchairs, indicating good consistency in GG item scores used in the measure score.

[4] 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.

4.3 Validity

4.3.1 Level(s) of Validity Testing Conducted *

Choose all that apply.

☒ Person– or encounter–level (i.e., data element) (e.g., sensitivity and specificity)
☒ Accountable entity–level (e.g., criterion validity)
☐ Not applicable/validity testing not conducted

4.3.1a Provide a rationale for why validity testing is not applicable/was not conducted

N/A

4.3.2 Type of accountable entity–level validity testing conducted *

Choose all that apply.

☒ Empirical validity testing at the accountable entity–level (e.g., criterion validity, construct validity, known groups analysis)
☐ Systematic assessment of face validity of the measure's performance score as an indicator of quality or resource use (i.e., the score is an accurate reflection of the effect of performance on quality or resource use and can distinguish good from poor performance).
☐ Not applicable/accountable entity–level validity testing not conducted

4.3.2a *[If a maintenance measure]* **Provide a rationale for why accountable entity–level validity testing was not conducted**

N/A

4.3.3 *[If validity testing was conducted]* Method(s) of Validity Testing *

For each level of testing conducted, describe the method(s) of validity testing and what each tests. Describe the steps (do not just name a method) and explain what was tested (e.g., accuracy of data elements compared with authoritative source, relationship to another measure as expected). What statistical analysis did you use? Describe proportion of missing data, how missing data was analyzed and/or excluded, and any sensitivity analysis conducted.

Note: Testing at the person or encounter level requires that all critical data elements be tested (not just agreement of one final overall computation for all patients). At a minimum, the numerator, denominator, and exclusions must be assessed and reported separately. For person or encounter level testing, prior evidence of validity of data elements for the data type specified in the measure (e.g., hospital claims) can be used as evidence for those data elements. Prior evidence could include published or unpublished testing that: includes the same data elements, uses the same data type (e.g., claims, chart abstraction), and is conducted on a sample as described above (i.e., representative, adequate numbers, and randomly selected, if possible).

For empirical accountable entity–level testing, the following should be included:

- *Narrative describing the hypothesized relationships*
- *Narrative describing why examining these relationships (e.g., correlating measures) would validate the measure*
- *Expected direction of the association*
- *Expected strength of the association*

We report testing results throughout this document at two levels: 1) data elements/scale and 2) Computed Quality Measure Score.

1. Critical Data Elements/Scale

Several studies have examined the validity of the data elements by examining the relationship between the items and length of stay, discharge to community rates and risk of falls.

2. Computed Quality Measure Score

Convergent Validity: To evaluate convergent validity of measure scores, we measured Spearman's rank correlation between the Cross-Setting Discharge Function Score for HH measure and other HH QRP measures. The analysis used CY 2023 data and only included data from HHAs with at least 20 stays.

To assess face validity of the Cross-Setting Discharge Function Score for HH measure, we convened two Technical Expert Panel (TEP) meetings (July 2021 and January 2022), as well as a Patient and Family Engagement Listening Session. 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. Additionally, panelists agreed that the Cross-Setting Discharge Function Score for HH measure adds value over the measures currently in-use in the SNF QRP (see Field 147 of the SNF MUC submission form).

The Patient and Family Engagement Listening Session demonstrated that the measure concept resonates with patients and caregivers. Participants understood and found important what self-care and mobility mean for patient outcomes. 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.

For the-use SNF QRP functional outcome measures—of which the Cross-Setting Discharge Function Score for HH is modeled—all missing item scores (i.e., Not Attempted, or NA, codes) are recoded to the code signifying the patient is completely dependent for an activity. However, TEP panelists agreed that NA codes may not always signify that a patient was dependent on a functional activity.[1] As a refinement, statistical imputation was implemented to estimate item scores for patients where a GG item was NA using models that adjust for patient clinical characteristics. We evaluated the empiric validity of our estimation methodology using the following analyses (see the Technical Report for full details).

1. We estimated admission and discharge scores for each GG item used in measure construction. To evaluate model fit of estimation models, we calculated C-statistics for each of the 22 estimation models. C-statistics ranged from 0.86-0.99, and the mean C-statistic was 0.96.

2. A bootstrapping method was used to measure bias and mean squared error (MSE) in the estimation method that uses statistical imputation compared to the recode approach used in the self-care and mobility functional outcome measures. Bias measures the average amount by which the estimated value differs from the true value. Bias is signed, with a positive amount meaning that the estimated 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. For the estimated values, average MSE was 1.44 at admission and 1.23 at discharge, and average bias was -0.22 at admission and -0.15 at discharge. For the recode approach, average MSE was 4.60 at admission and 13.30 at discharge, and average bias was -0.54 at admission and -0.70 at discharge. This result indicates that the estimation approach produced less biased, more precise estimates for missing item scores.

3. We calculated the difference in discharge function between episodes that have bona fide item scores at admission and stays with NA codes at admission where we estimate the item score. This difference provides a metric of how accurately estimated item scores reflect true patient function. For all 11 items, the difference was lower than if these ANAs were recoded to the most dependent level of functional status. This result indicates that statistical estimation produced more accurate results.

4.3.4 *[If validity testing was conducted]* Validity Testing Results *

Provide the statistical results from validity testing for each level and type of validity testing conducted.

A Technical Expert Panel provided feedback on the Cross-Setting Discharge Function for HH measure representing face validity.

1. **Expert Consensus on Cross-Setting Discharge Function Score for HH**
   - The discharge function measure was reviewed and supported by a multi-disciplinary panel of experts, including persons with lived experience.
   - **Evidence**: The panelists favored reporting discharge function measures due to their ability to reflect patient recovery at discharge. They preferred reporting discharge function rather than change in function measures because it better captures patient status at the point of leaving the provider.

"Panelists from the July 2021 TEP favored Discharge Function Score measures over Change in Function Score measures and recommended moving forward with Discharge Function Score for the cross-setting measure."
Source: PAC Function TEP Summary Report – July 2021 and PAC Function TEP Summary Report – January 2022.

2. **Robust Risk Adjustment for Fair Comparisons**
   - The measure uses a robust **risk adjustment** methodology, which supports fair comparisons of measure results across providers by accounting for differences in patient age, clinical characteristics and comorbidities.
   - **Evidence**: This ensures that providers are compared on a level playing field, taking into account the complexity of patients treated at each provider.

"Calculate expected Discharge Function Mobility Score for each eligible stay using risk adjustment coefficients, including demographics, health characteristics, and admission function score."
Source: PAC Function TEP Summary Report – January 2022.

3. **Handling of Activities Not Attempted codes**
   - The discharge function measure incorporates **statistical imputation** to address that not all patients can complete each of the functional activities and are thus coded using the **Activities Not Attempted codes. This supports measure** validity even when certain activities cannot be completed during the patient's stay.

- **Evidence**: The TEP strongly favored using **statistical imputation** over simply coding missing data as "dependent," ensuring that the discharge function measure more accurately reflects the patient's true capabilities.

"Panelists tended to favor statistical imputation with continued refinement to improve cross-setting performance. Panelists agreed that the current recode could be improved upon."
Source: PAC Function TEP Summary Report – July 2021 and PAC Function TEP Summary Report – January 2022.

4. **Alignment with Patient-Centered Outcomes**
   - The discharge function measure is designed to reflect patient-centered goals, focusing on the safe and functional transition of patients back to their community or home setting.
   - **Evidence**: Functional outcomes at discharge are aligned with patient goals of regaining independence, which is a key measure of quality in post-acute care.

"The discharge function score is designed to reflect the ability of post-acute care providers to successfully rehabilitate patients, ensuring they regain functional independence at discharge and beyond."
Source: PAC Function TEP Summary Report – January 2022.

5. **Interested Parties Engaged and Broad Support**
   - The measure was reviewed by a diverse group of interested parties with broad support and clinical relevance across different care settings.
   - **Evidence**: Clinicians, policy experts, and performance measurement specialists contributed their feedback, ensuring that the measure is relevant and usable across different PAC settings.

"The PAC QRP Functions TEP comprised 15 stakeholders with diverse perspectives and areas of expertise, including clinical, policy and program, measures development, and technical expertise."
Source: PAC Function TEP Summary Report – January 2022.

Measure validity was assessed by comparing the DC Function measure with other quality measures in the HH Quality Reporting Program using Spearman (rank) correlations between provider's performance scores.

As expected, this measure demonstrated positive correlation with the Discharge to Community measure (0.26), which was significant ($p<0.01$). Correlation was also positive with Improvement in Ambulation (0.31), Improvement in Bed Transfer (0.45), Improvement in Bathing (0.32), Improvement in Dyspnea (0.35), and Improvement in Oral Medication Management (0.28).

Exhibit 9. Correlation Between the Discharge Function Score for HH Measure and Other HHA QRP Measures (Publicly Reportable Providers)

HH QRP Measures	N Providers	Other HH QRP Measure		Discharge Function		Correlation	
		Mean	SD	Mean	SD	Spearman Correlation	P Value
Assessment Based Measures							
Improvement in Ambulation-Locomotion	8,091	81.91	14.76	57.35	19.55	0.31	0.000
Improvement in Bathing	8,091	84.41	14.46	57.35	19.55	0.32	0.000

Improvement in Bed Transferring	8,091	81.79	15.52	57.35	19.55	0.45	0.000
Improvement in Dyspnea	8,088	83.33	16.54	57.35	19.54	0.35	0.000
Falls with Major Injury	8,091	1.06	1.15	57.35	19.55	0.08	0.000
Improvement in Management of Oral Medications	8,091	78.77	16.84	57.35	19.55	0.28	0.000
Claims Based Measures							
PPR	7,754	3.88	0.37	57.64	19.21	-0.02	0.116
DTC	7,879	76.46	12.17	57.47	19.49	0.26	0.000
MSPB	7,876	0.97	0.15	57.46	19.48	0.08	0.000
PPH	7,847	10.15	2.76	57.45	19.48	-0.17	0.000

4.3.4a *[If validity testing was conducted]* Attach Additional Validity Testing Results

If needed, you may attach additional validity testing results here. Please ensure all attachments are 508 compliant, all tables and figures are labeled with alternative text, as appropriate. Please clearly refer to any results within your attachment within the relevant text fields of this measure submission form.

One file only; 256 MB limit; Allowed types: .zip, .pdf, .docx, .xls, .xlsx

4.3.5 *[If validity testing was conducted]* Interpretation of Validity Results *

Provide your interpretation of the results in terms of demonstrating validity for each level and type of validity testing conducted. How do the results support an inference of validity for the measure? For accountable entity–level testing, discuss how the results relate to the hypothesis? If the results are not what were expected, why?

1. Critical Data Elements: We reviewed results from several published studies that examined the validity of the function items.
2. Computed Quality Measure Score

Convergent Validity. First, as expected, scores for the Cross-Setting Discharge Function Score for HH measure correlated well but not perfectly with other cross-setting measures including Improvement in Ambulation-Locomotion (0.31) and Improvement in Bed Transferring (0.45). This is expected because the HH QRP functional improvement measures measure whether the HHA improved patient function, while the Cross-Setting Discharge Function Score for HH measures whether patient function exceeds expectations at discharge. Second, the Cross-Setting Discharge Function Score for HH 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. The Cross-Setting Discharge Function Score for HH is associated with discharge to community, and this measure demonstrated the expected positive correlation with the Discharge to Community measure (0.26). Additionally, it was negatively correlated with the Potentially Preventable Readmissions (PPR) within 30-Days Post-Discharge measure (-0.02)

and the Potentially Preventable Hospitalization measure (-0.17). The measure had a weak positive correlation with Medicare Spending Per Beneficiary (0.08). All correlation coefficients were significant (p<0.01) with the exception of the PPR measure.

4.4 Risk Adjustment

4.4.1 Methods Used to Address Risk Factors *

*What methods or approaches were used to explore the effects of risk factors on this measure? (**Note:** If you tested for the effects of risk factors and ultimately determined that risk adjustment or stratification was not warranted, please select the method(s) used and provide details of the testing and your rationale in 4.4.2 through 4.4.6; the measure's ultimate status will be reported in 4.4.7). Choose all that apply.*

☒ Statistical risk adjustment model with risk factors
☐ Stratification by risk factor category
☐ Other

4.4.1a Describe other method(s) used

N/A

☐ No risk adjustment or stratification.

4.4.1b *[If Measure Type is outcome, cost/resource, or PRO–PM]*
Provide a rationale for why there is no need to address differences in patient characteristics (i.e., case mix) to achieve fair comparisons across measured entities for your outcome or resource measure.

N/A

4.4.2. *[If risk factors are addressed by any method (4.4.1)]* Conceptual Model Rationale *

Explain the rationale for the risk approach, including reasons for risk adjustment and/or stratification. Describe the sources that inform the conceptual model, e.g., scientific literature, unpublished findings, TEP. Consider age, gender, race, ethnicity, urbanicity/rurality, Medicare/Medicaid dual eligibility status, indices of social vulnerability (e.g., Centers for Disease Control and Prevention Social Vulnerability Index), and markers of functional status–related risk (e.g., cognitive or physical function) in the conceptual model, using evidence to support the model, with references. If risk factors (e.g., social, functional status–related, clinical) are included in the conceptual model but data are not available for all factors, describe any potential bias, as a result of not including the risk factor(s) in the final risk adjustment model or stratification. Address the validity of the measure in light of this bias.'

The rationale for risk adjustment is to account for differences in patient populations. By risk adjusting, the performance measure assesses providers based on their quality of care and not the underlying health of the population. Providers are not penalized for serving patients with greater clinical need, and fair comparisons can be made across providers.

The performance measure is cross-setting, calculated for inpatient rehabilitation facilities (IRF), skilled nursing facilities (SNF), long-term care hospitals (LTCH), and home health (HH). The

development team sought to align risk factors across settings as much as possible. The team presented to a TEP an overview of the availability of clinically meaningful risk factors in each setting. TEP members supported setting-specific parameters for risk adjustment since there are different data points available as well as clinical considerations for each setting.

The development team also presented to the TEP the conceptual model shown below in 4.4.2a. TEP members agreed that the conceptual model presented does represent the salient points about the relationship between social risk factors (SRFs), patient functional outcomes, and provider quality. TEP members provided examples of ways in which providers are able to, and should be expected to, mitigate the influence of SRFs on patient outcomes.

TEP members supported further analysis to understand the effect of measurable SRFs. Specifically, the TEP cited the following as potential measurable SRFs that can impact functional outcomes: dual enrollment, ADI, and race/ethnicity (although noting that these are impacted by provider bias).

Below are the currently measurable SRFs included in risk adjustment testing, but not used in the final risk adjustment model. Health-related social needs items are not yet available cross-setting, but can be tested for inclusion in the future.

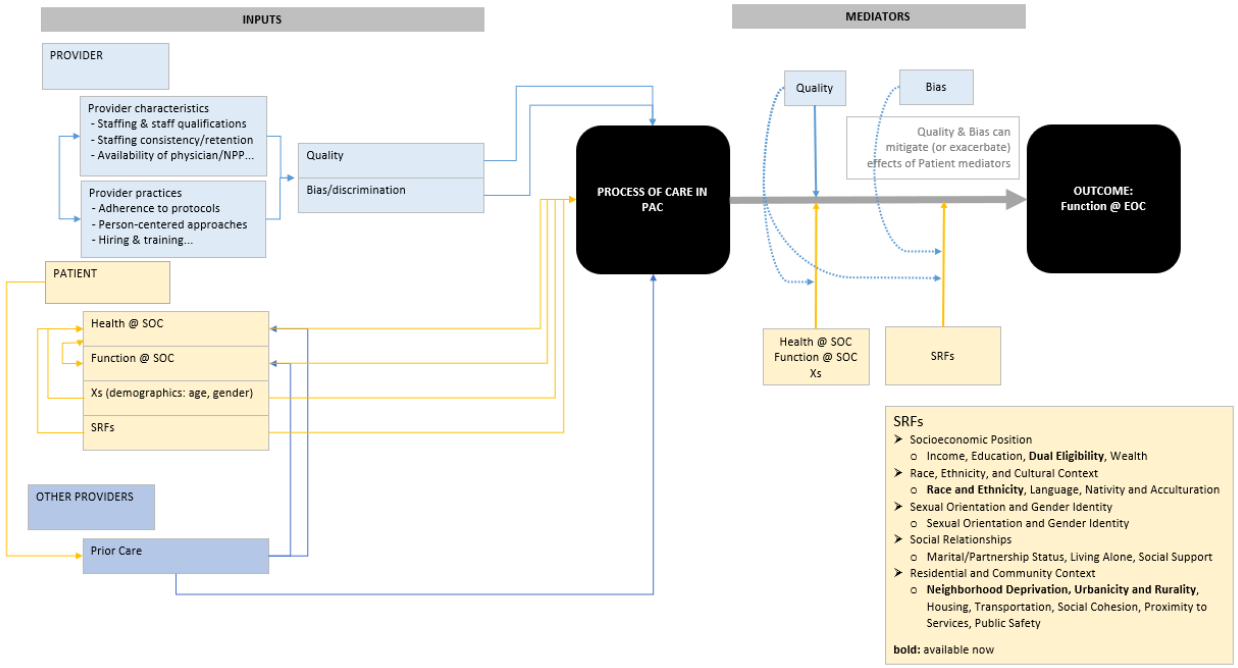
Exhibit 10. Social Risk Factors (SRFs) Included in Risk Adjustment Testing

Dual Enrollment
Medicare (reference group)
Dual
Medicaid
Neither Medicare nor Medicaid
Unknown Payer
Race
American Indian or Alaska Native
Asian
Black
Hispanic or Latino
Multiple Race
Native Hawaiian or Other Pacific Islander
No R/E Information Available
White (reference group)
Area Deprivation Index (ADI)
ADI (≥85)
ADI Missing

4.4.2a *[If risk factors are addressed by any method (4.4.1)]* **Attach Conceptual Model** *

Attach a figure of the conceptual model that illustrates the hypothesized pathway between the social and/or functional status–related risk factors, patient clinical factors, quality of care, and the measured outcome.

One file only; 256 MB limit; Allowed types: .pdf, .jpg, .png, .zip



4.4.3 *[If risk factors are addressed by any method (4.4.1)]* Risk Factor Characteristics Across Measured Entities *

Provide descriptive statistics showing how the risk variables identified from the conceptual model are distributed across the measured entities. Indicate which risk factors were tested in the risk adjustment model and which were tested for stratifying the measure, as applicable.

The table attached for Section 4.4.4a presents information for each risk factor covariate in the final model plus the additional SRF risk factors considered but not used in the final risk adjustment model.

4.4.4 *[If risk factors are addressed by any method (4.4.1)]* Risk Adjustment Modeling and/or Stratification Results *

Describe the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model and/or stratification, as applicable. Clearly indicate the risk factors included in the final risk model and/or used in the final stratification approach.

Cross-Setting Discharge Function Score for HH is a cross-setting performance measure calculated for IRF, SNF, LTCH, and HH. Because different data elements are collected across the assessment instruments for each setting, the development team aligned clinically meaningful covariates as much as possible.

The development team then presented to a TEP an overview of the data availability in each setting, shown below, and solicited feedback on which covariates should be included in the cross-setting measure risk adjustment model.

Exhibit 11. Covariate Availability Across Settings of Care for Cross-Setting Discharge Function Score Measure

Covariate	LTCH	IRF	SNF	HH
Age	✓	✓	✓	✓
Admission Mobility Score	✓	✓	✓	✓
Primary Medical Condition Category (PMCC)	✓	✓	✓	
Interaction of Admission Mobility Score and PMCC	✓	✓	✓	
Prior Function/Device Use	✓	✓	✓	✓
Pressure Ulcers	✓	✓	✓	✓
Cognitive Function		✓	✓	✓
Communication Impairment	✓	✓	✓	
Incontinence	✓	✓	✓	✓
Falls		✓	✓	✓
Nutritional Approach	✓	✓	✓	
Comorbidities	✓	✓	✓	✓
Ventilation Status	✓			
Availability of Assistance				✓
Living Arrangements				✓
Entry Pattern				✓

The TEP members expressed support for setting-specific models since there are different data points available as well as different clinical considerations for each setting. The panelists suggested additional risk adjustors to consider, including Prior living site; Prior hospitalization; Chronic conditions; Obesity; Severity of health condition(s); Low BMI; Pain; Wound infection; Transportation; and Health literacy.

Below is a listing of the covariate groups included in the final risk adjustment model for HH. 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, 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 estimated scores (Please see Steps 1 – 2 in 1.18). The walking and wheelchair activities are used in the same manner as for the discharge score (Step 2 in Section 1.18). 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 ≤ BMI ≤ 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.

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.

The risk adjustment model is an ordinary least squares (OLS) linear regression. It estimates the relationship between discharge function score and the set of risk adjustors. The risk adjustment model is run on all HHA quality 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 quality episode at SOC/ROC.

The risk adjustment model is written as

$$Expected\ Discharge\ Function\ Score = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n$$

where $x_1 - x_n$ are the risk adjustors.

We also tested three SRFs of interest that ultimately were not included in the final risk adjustment model:

1. Medicare vs. dually enrolled (patient is dually enrolled at any time during the quality episode)
2. Race/ethnicity
3. ADI

We constructed alternative risk adjustment models that included additional covariates for payer, race/ethnicity, and ADI to consider these SRFs for inclusion in the risk adjustment model.

4.4.4a *[If risk factors are addressed by any method (4.4.1)]* **Attach Risk Adjustment Modeling and/or Stratification Specifications** *

Provide detailed risk adjustment model and/or stratification specifications, including the method(s), risk factor data sources, and equations, as applicable Please list all risk factors in your conceptual model, clearly indicating which factors were available/tested

and which (if any) were retained in final model and/or stratification plan. Also include the data source, code with descriptor, and coefficient for each risk factor in the final risk adjustment model or stratification plan, as appropriate.
One file only; 256 MB limit; Allowed types: .xls; .xlsx; .csv

4.4.5 *[If 4.4.1 includes "Statistical risk adjustment model with risk factors"]* **Calibration and Discrimination** *

Describe the approach and results of calibration and discrimination testing. Describe any over– or under–prediction of the model for important subgroups.

A well-calibrated model demonstrates good predictive ability to distinguish high-risk from low-risk patients. To assess risk adjustment model calibration, we calculated the ratio of observed-to-predicted discharge function score across eligible stays by decile of predicted discharge function score (risk).

The average ratio 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. Below are the ratios overall and by decile.

We analyzed model fit 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 0.48, which suggests good model discrimination.

4.4.5a *[If 4.4.1 includes "Statistical risk adjustment model with risk factors"]* **Attach Calibration and Discrimination Testing Results** *

Attach results of calibration and discrimination testing.

One file only; 256 MB limit; Allowed types: .pdf; .zip

4.4.6. *[If risk factors are addressed by any method (4.4.1)]* **Interpretation of Risk Factor Findings** *

Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix). Clearly describe the rationale for why each risk factor tested WAS or WAS NOT included in the final model. Describe what the results mean, including what is normally expected in relation to the test conducted.

Risk factors were chosen based on clinical relevance to Cross-Setting Discharge Function Score for HH performance. Risk factors were recommended by clinician members of the measure development team and by the TEP. The final risk adjustment model has an adjusted R-squared of 0.48.

We find that across the alternative risk adjustment models considered, the SRF covariates are significant but small, and have little to no impact on model fit. The details of the alternative risk adjustment models are shown in the attachment for Section 4.4.4a. The attached file presents the model results for the final risk adjustment model and the alternative risk adjustment model with additional SRF covariates.

While we considered these SRFs for inclusion in the risk adjustment model, we ultimately decided against such inclusion, primarily for conceptual reasons. Including these SRFs in risk

adjustment models runs the risk of adjusting for factors that providers could control and should improve on – like active/unconscious bias against particular patient populations (e.g., more actively accommodating different levels of health literacy, better access to interpreter services for people who are not native English speakers, better accommodations for disabled patients). It effectively lowers the expected outcomes for high-SRF patients, making expectations easier to meet, without improving the actual outcomes or underlying treatments. Further, when measures are stratified by such SRFs (enabling identification of gaps in provider quality between, for example, dually and non-dually enrolled patients, as is done in confidential feedback reports to providers), adjusting for dual eligibility as a risk factor may diminish CMS's ability to make such stratified information clear and useful to providers. Finally, assessment items released since measure development allow for the possibility of more refined measurement of social determinants of health (e.g., health literacy, transportation). These alternatives can be tested for future revisions of the Cross-Setting Discharge Function Score for HH measure.

4.4.7 *[If risk factors are addressed by any method (4.4.1)]* **Final Approach to Address Risk Factors** *

*After testing, what methods or approaches were ultimately used to control for the effects of risk factors? (**Note:** the final approach should be supported by the testing and the rationale provided in 4.4.2–4.4.6). Choose all that apply.*

☒ Statistical risk adjustment model with risk factors
☐ Stratification by risk factor category
☐ Other

4.4.1a Describe other method(s) used

N/A

☐ No risk adjustment or stratification.

Section 5. Equity

5.1 Contributions Towards Advancing Health Equity (*optional*).

Describe how this measure contributes to efforts to advance health equity Provide a description of your methodology and approach to empirical testing of differences in performance scores across multiple socio–contextual variables (e.g., race, ethnicity, urbanicity/rurality, socio–economic status, gender, gender identity, sexual orientation, age). Provide an interpretation of the results, including interpretation of any identified differences and consideration of negative impact or unintended consequences on subgroups.

The measure at hand provides a means for assessing the impact of provider performance on patients who experience social risk factors (SRF) to a greater degree than those who have fewer or less acute SRFs. For example, dual-eligible patients tend to experience worse socioeconomic circumstances than other patients. These circumstances can negatively impact health outcomes. Some of the disparity in outcomes between dual and non-dual patients can be explained through differences in prevalence of clinical conditions addressed through risk adjustment. However, even after risk adjustment, dual patients fare worse, on average, than non-duals for all settings. One contributing factor could be that there are socioeconomic drivers of health disparities in dual patients beyond what is captured through risk adjustment. This raises the concern that providers who serve these populations are unduly penalized in quality measurement when dual-eligibility is not included in the risk adjustment model.

We tested three SRFs of interest:

1. Medicare vs. dually enrolled (patient is dually enrolled at any time during the quality episode)
2. Race/ethnicity
3. ADI

We used several approaches to test differences in performance scores across multiple SRFs and to consider some SRFs for inclusion in the risk adjustment model. First, we constructed alternative risk adjustment models that included additional covariates for payer, race/ethnicity, and ADI, and examined the impact on provider performance.

We find that across most of the alternative risk adjustment models considered, the SRF covariates are significant but small, and have little to no impact on model fit. The details of the alternative risk adjustment models are shown in the attachment for Section 4.4.4a.

Second, we stratified the performance scores by SRFs. Using the current model, we calculated provider scores for patients with and without SRFs and grouped HHAs into quintiles based on their proportion of Black/non-White, dual, and dual and high ADI patients. We then examined whether performance declines with the proportion of patients with SRFs, and whether this impacts patients both with and without SRFs.

Across home health agencies, we compared Cross-Setting Discharge Function Score for HH CY 2023 performance by subgroups of agencies based on the percentage of patients who are Black, Non-White, Medicaid or Dual-eligible, and Dual-eligible or living in a neighborhood with ADI ≥ 85. To be more specific, we defined subgroups of agencies based on quintiles of the percentage of patients within the agency who have the SRF. For race and ethnicity

characteristics, we used the M0140: Race/Ethnicity OASIS item to identify patient's race/ethnicity as Black or Non-White.

When examining performance by proportion of high-SRF patients served, Cross-Setting Discharge Function Scores for HH for all patients (both high- and low-SRF) decrease as proportion of high-SRF patients served by the provider increases (see Exhibit 12 below). On average, scores for low-SRF patients tend to be slightly higher than for high-SRF patients. This relationship holds true within each quintile category of providers based on proportion of high-SRF patients served. We find a gap of a few percentage points in mean final score between SRF patient populations, and that the mean final scores decrease with each quintile. However, we also find that in most cases the ratio of observed/predicted is slightly lower for the SRF population.

Exhibit 12. Episode-Level Cross-Setting Discharge Function Scores for HH Stratified by Proportion of Social Risk Factors Served by Provider (Publicly Reportable Providers with 10+ Stays in Each Population)

Characteristic (Population 1)	Quintile of Provider-Level % of Characteristic	N Stays	N Providers	Final Score		
				Overall	Pop. 1	Pop. 2
					Black	White
Black	Overall	4,346,014	3,789	63.2%	60.3%	64.7%
Black	Quintile 1	29,319	8	69.6%	64.1%	70.0%
	Quintile 2	693,249	424	65.1%	61.8%	66.0%
	Quintile 3	1,087,538	918	64.1%	60.9%	65.2%
	Quintile 4	1,299,991	1,161	63.6%	61.4%	64.7%
	Quintile 5	1,235,917	1,278	60.9%	59.6%	62.7%
					Non-White	White
Non-White	Overall	5,109,294	7,245	62.9%	60.4%	64.4%
Non-White	Quintile 1	1,076,564	1,429	67.5%	66.4%	67.8%
	Quintile 2	1,548,186	1,564	64.8%	63.5%	65.3%
	Quintile 3	1,177,548	1,588	61.9%	60.6%	62.7%
	Quintile 4	933,895	1,594	59.1%	58.1%	60.1%
	Quintile 5	373,101	1,070	54.5%	53.7%	56.5%
					Dual	Medicare
Medicaid or Dual	Overall	5,031,926	6,602	63.0%	59.6%	64.2%
Medicaid or Dual	Quintile 1	1,158,649	1,249	63.9%	60.8%	64.3%
	Quintile 2	1,409,179	1,521	63.4%	60.7%	64.0%
	Quintile 3	1,524,037	1,548	65.1%	63.2%	65.9%
	Quintile 4	777,057	1,567	59.4%	57.4%	61.1%
	Quintile 5	163,004	717	51.2%	50.7%	53.3%

					Dual or ADI ≥ 85	Medicare or ADI<85
Dual or ADI ≥ 85	Overall	4,982,522	6,397	63.2%	59.6%	65.0%
Dual or ADI ≥ 85	Quintile 1	1,280,585	1,252	64.1%	60.6%	65.7%
	Quintile 2	1,573,237	1,500	64.2%	61.3%	65.4%
	Quintile 3	1,319,279	1,510	64.7%	62.7%	65.9%
	Quintile 4	662,010	1,452	58.7%	57.1%	60.7%
	Quintile 5	147,411	683	50.0%	49.3%	50.7%

We also examined the distribution of within-provider performance gaps by SRFs (e.g., duals vs. Medicare) to demonstrate the range of gaps. Using the dually eligible population as an example, we find considerable variation in within-provider performance gaps between dually and non-dually enrolled patients. A substantial number of providers had either no performance gap between dually and non-dually enrolled patients or had better performance in their dually eligible patients, indicating that there are providers who successfully produce comparable or better outcomes for dual patients. The proportion of providers who met or exceeded expectations for an equivalent or larger share of their dual patients is 36.1%. This suggests that there's room for providers to improve care for dually enrolled patients.

Section 6. Use & Usability

[NOTE: Current/Planned Use and Program Details (Section 6.1, items 6.1.1–6.1.4) are now entered in the ITS and can be edited in the FMS]

6.2 Usability

6.2.1 Actions of Measured Entities to Improve Performance *

What are the actions measured entities must take to improve performance on this measure? How difficult are those actions to achieve and how can measured entities overcome those difficulties?

All home health agencies with at least 20 qualifying quality episodes of care receive quarterly measure reports on all their publicly reported measures. In addition, providers can run on-demand, confidential reports showing individual measure results and national averages, through CMS' CASPER system. There is an email box that HHAs may submit questions to as well as a website on which the latest measure updates are posted. The OASIS Guidance Manual describes the OASIS-based reports that are available, report use(s) and provides guidance about OASIS and quality improvement. Home health agencies make use of these reports to monitor and improve the quality of care.

6.2.2 *[If maintenance review OR Current Status = Yes (6.1.1)]* Feedback on Measure Performance *

Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how you obtained feedback.

Home health agencies receive quarterly measure reports on all their measures. There is an email box that HHAs may submit questions to as well as a website on which the latest measure updates are posted. Because of the changes made to the OASIS in OASIS E (effective January 1, 2023), risk models for publicly reported outcome measures have been updated. CMS makes available information about risk models and covariates on its website.

6.2.3 *[If maintenance review OR Current Status = Yes (6.1.1)]* Consideration of Measure Feedback *

Describe how you considered the feedback when developing or revising the measure specifications or implementation, including whether you modified the measure and why or why not.

No measure specifications changes requested or made.

6.2.4 *[If maintenance review OR Current Status = Yes (6.1.1)]* Progress on Improvement *

Discuss any progress on improvement (trends in performance results, including performance

across sub–populations if available, number and percentage of people receiving high–quality health care, geographic area, number and percentage of accountable entities and patients included). If use of the measure demonstrated no improvement, provide an explanation.

This measure is too new to provide an assessment of impacts on improvement.

6.2.5 *[If maintenance review OR Current Status = Yes (6.1.1)]* **Unexpected Findings** *
Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

None

Section 7. Supplemental Attachment

7.1 Supplemental Attachment

If needed, you may attach additional measure information here. Please ensure that all included files are 508 compliant, including labeling all tables and figures with alternative text, as appropriate. Clearly label all components of the attachment with the field number(s) its contents refer to, and likewise, clearly refer to any results in this attachment within the relevant text fields of the FMS.

One file only; 256 MB limit; Allowed file types: .zip; .pdf;.docx; .xlsx

Appendix A: Imputation

In a preliminary series of steps in the process of calculating DC Function, GG items at start of care or resumption of care (SOC/ROC) and at discharge that have an Activity Not Attempted (ANA) code of 07, 09, 10, or 88, a dash (-), or a skip (^) recorded (hereafter referred to as NA) are estimated using statistical imputation. The estimation models include the predictors used in risk adjustment and covariates for scores on other GG items. Notably, the estimation models use all GG items available in HH to estimate the ANA scores for the subset of GG activities used for the DC Function numerator. After estimation then, in the second phase, the calculation of DC Function continues.

The steps below describe how to estimate a single item at SOC/ROC and then describe the relevant modifications for estimating the item at discharge for the other items.

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

Step 2: For each HH quality episode, determine whether to use walking or wheelchair items in the estimation model.

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 items.

Otherwise, use walking items.

Step 3: Create variables for the estimation model reflecting how each item except Eating (GG0130A) was scored at SOC/ROC. 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, the second reflects 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 activity was skipped (g^{**}).

$$Function\ items: G \in \{g_2, \ldots, g_{10}\}$$

$$g' = [g, g^*, g^{**}]$$

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

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

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

$$Function\ items\ with\ NA\ indicators: G' \in \{g'_2, \ldots, g'_{10}\}$$

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 estimation method: clinical covariates (C) and function activities with NA indicators (G') constructed in Step 3.

$$Clinical\ items := C \in \{c_1, \ldots, c_k\}$$

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

The model we estimate for g_1, GG0130A, is

$$z_i = C_i\beta + G_i'\phi + \varepsilon_i$$

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

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 estimated value of g_i (rounded to four decimal places) as

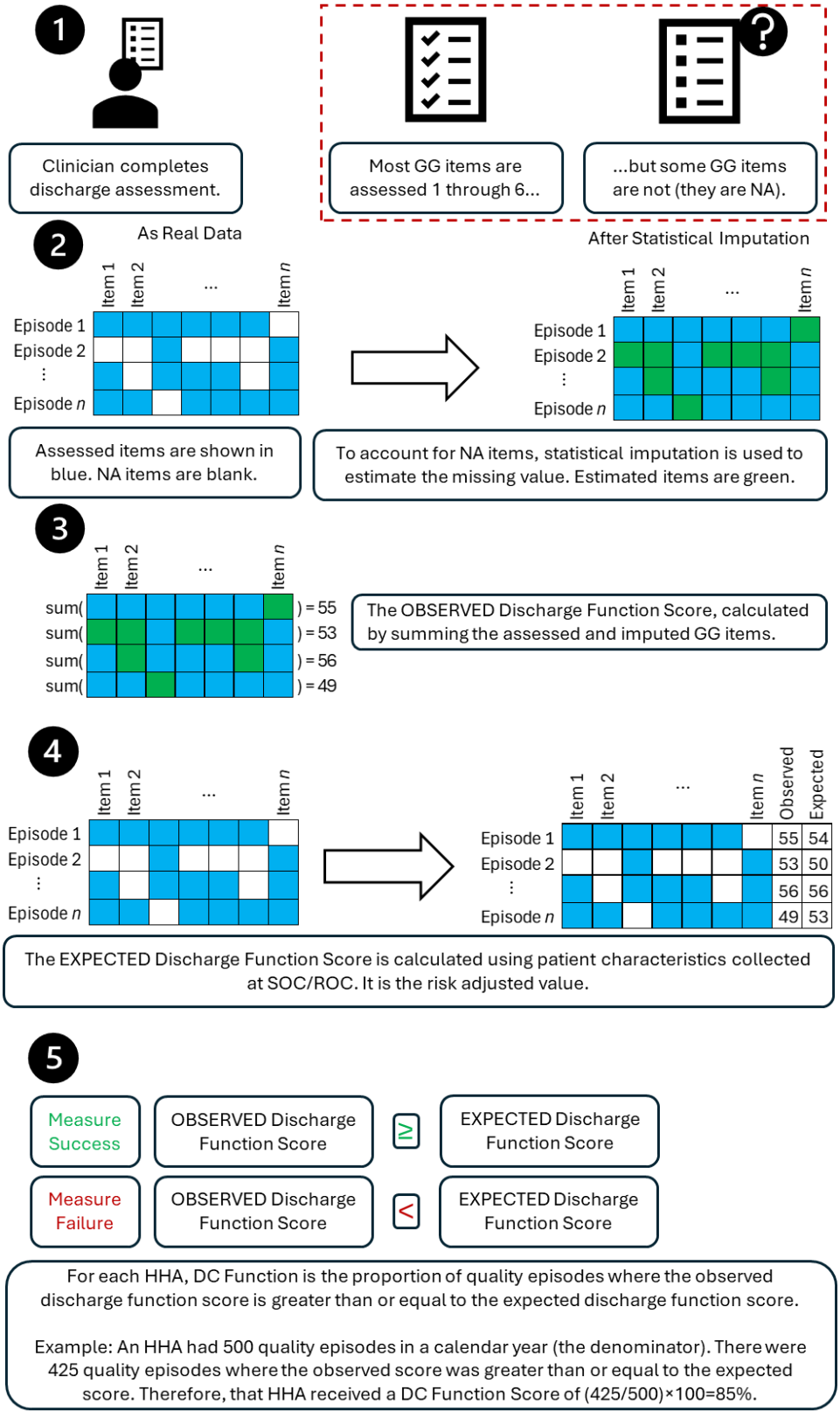
$$\widehat{g_i} = \Pr(z_i \leq \alpha_1) + 2 * \Pr(\alpha_1 < z_i \leq \alpha_2) + 3 * \Pr(\alpha_2 < z_i \leq \alpha_3) + 4 * \Pr(\alpha_3 < z_i \leq \alpha_4) + 5 * \Pr(\alpha_4 < z_i \leq \alpha_5) + 6 * \Pr(z_i > \alpha_5)$$

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 item included in the observed discharge function score, 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 first set of conditions in Step 2. For Walk 10 Feet (GG0170I) and Walk 50 Feet with 2 Turns (GG0170J), use only the sample of quality episodes that satisfies the second set of conditions in Step 2.

The steps above are summarized in the following Exhibit A-1.

Exhibit A-1. Summary of steps for statistical imputation of GG items not assessed (NA)



Appendix B: Measuring bias and mean squared error (MSE) in the imputation method

A bootstrapping method was used to measure bias and mean squared error (MSE) in the estimation method for statistical imputation compared to the recode approach used in the self-care and mobility functional outcome measures. Bias measures the average amount by which the estimated value differs from the true value. Bias is signed, with a positive amount meaning that the estimated 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 how similar estimated values were to the true item score. For each bootstrap, episodes with complete item data were sampled using stratified random sampling. Two copies were made of this sample. The first copy was the original with known item scores. Missing item scores were imposed on the second copy, and now-missing item scores were estimated using both statistical modeling and the recode approach. Item scores estimated through each approach were compared to the known item scores from the first copy. The MSE and bias statistics were calculated as averages across bootstraps. For statistical estimation, average MSE was 1.44 at admission and 1.23 at discharge, and average bias was -0.22 at admission and -0.15 at discharge. For the recode approach, average MSE was 4.60 at admission and 13.30 at discharge, and average bias was -0.54 at admission and -0.70 at discharge. This result indicates that statistical estimation produced less biased, more precise estimates for missing item scores.