1.18 Calculation of Measure Score

Appendix A: Imputation

In a preliminary series of steps in the process of calculating DC Function, GG items at admission 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 the post-acute care setting (SNFs, IRFs or LTCHs) 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 admission and then describe the relevant modifications for estimating the item at discharge for the other items.

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

<u>Step 2:</u> For each stay, determine whether to use walking or wheelchair items in the estimation model.

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

Otherwise, use walking items.

<u>Step 3:</u> Create variables for the estimation model reflecting how each item except Eating (GG0130A) was scored at admission. GG item scores are described as independent variables (i.e., on the right-hand side) by three variables, collectively referred to as g'. The first reflects a score of 1 – 6, 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^{**}).

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

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

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

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

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

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

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.

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

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

The model we estimate for g_1, GG0130A, is

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

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

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

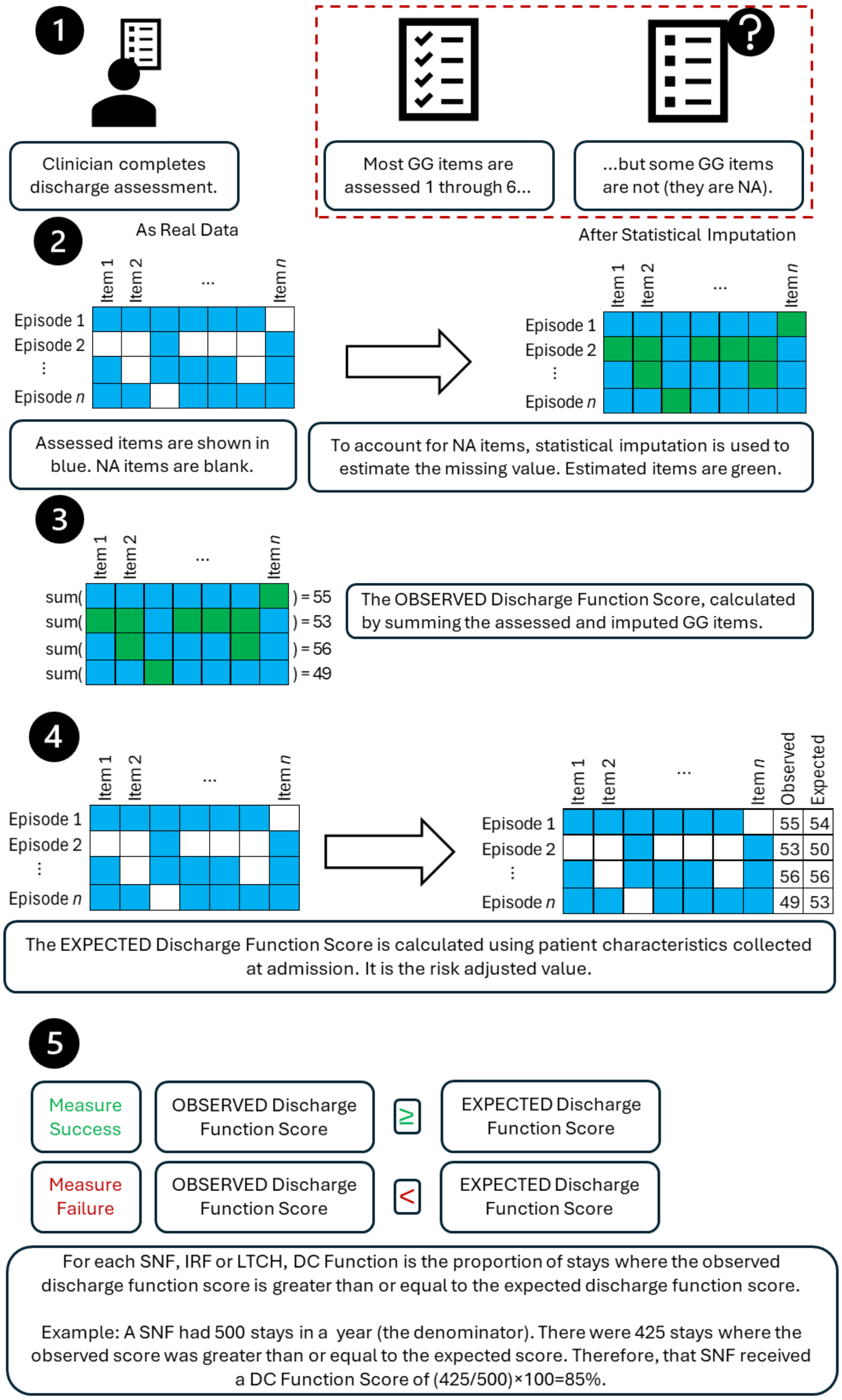
$$\hat{g_i} = \Pr(z_i \le \alpha_1) + 2 * \Pr(\alpha_1 < z_i \le \alpha_2) + 3 * \Pr(\alpha_2 < z_i \le \alpha_3) + 4 * \Pr(\alpha_3 < z_i \le \alpha_4) + 5 * \Pr(\alpha_4 < z_i \le \alpha_5) + 6 * \Pr(z_i > \alpha_5)$$

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 stays 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 stays that satisfies the second set of conditions in Step 2.

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

Exhibit A-1:



The estimation coefficients and thresholds are available publicly in the "download" section for each setting at:

- Skilled Nursing Facility (SNF) Quality Reporting Program (QRP) Measures and Technical Information | CMS.
    - Direct link to file: https://www.cms.gov/files/document/imputation-appendix-file-snf-effective-10-01-2024.xlsx
- Inpatient Rehabilitation Facility Reporting Measures Information
    - Direct link to file: https://www.cms.gov/files/document/imputation-appendix-file-irf-effective-10-01-2024.xlsx
- Long-Term Care Hospital (LTCH) Quality Reporting Program (QRP) Measures Information | CMS
    - Direct link to file: https://www.cms.gov/files/document/imputation-appendix-file-ltch-effective-10-01-2024.xlsx

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, which recodes all NAs to 1. 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.

In SNFs, statistical estimation resulted in lower levels of bias (-0.21 at admission; -0.17 at discharge) and MSE (1.71 at admission; 1.41 at discharge) compared to the bias (-1.24 at admission; -0.72 at discharge) and MSE (5.05 at admission; 4.18 at discharge) produced from the recode approach, which supports the validity of the statistical estimation method for this setting.

In IRFs, statistical estimation also resulted in lower levels of bias (-0.39 at admission; -0.07 at discharge) and MSE (2.17 at admission; 0.50 at discharge) compared to the bias (-1.43 at admission; -0.51 at discharge) and MSE (6.99 at admission; 2.58 at discharge) produced from the recode approach, which supports the validity of the statistical estimation method in this setting.

In LTCHs, statistical estimation resulted in lower levels of bias (-0.02 at admission; -0.24 at discharge) and MSE (5.98 at admission; 3.12 at discharge) compared to the bias (-2.73 at admission; -1.54 at discharge) and MSE (20.21 at admission; 9.80 at discharge) produced from the recode, which supports the validity of the statistical estimation method in LTCHs. This result indicates that statistical estimation produced less biased, more precise estimates for missing item scores across settings.