NATIONAL QUALITY FORUM

Memo

June 11, 2019

To: Patient Experience and Function (PEF) Committee

From: PEF Project Team

Re: Competing Measures

Background

During the 2015 Person and Family Centered Care (PFCC) project, two pairs of measures specified for use in Inpatient Rehabilitation Facilities (IRFs) were identified as competing and required additional consideration from the Consensus Standards Approval Committee (CSAC) and NQF Board of Directors (Board). These pairs of measures received considerable discussion and public comment, including review and deliberations by the PFCC Standing Committee, the CSAC, and the Board. Ultimately, the PFCC Standing Committee could not come to a best-in-class decision, and all four of the measures were endorsed. However, the Board strongly requested that the Committee pick a best-in-class measure at the next maintenance review.

These four measures have now been resubmitted for maintenance of endorsement.

This memo contains:

- NQF's guidance on competing measures;
- The summary of the 2015 measure review process and outcomes;
- Comments received on the competing measures during the pre-meeting comment period;
- Action steps for the Committee;
- Additional information submitted by the developers at the Committee's request during the orientation call
    - Memo from UDSMR
    - Memos from CMS & RTI
- A side-by-side comparison of each of the pairs of competing measures
- The relevant sections of the 2015 report detailing the full discussion and evaluation of each of the four measures
- The 2015 memo to the Board detailing the process and the conditions for endorsement

2019 Submissions

In the spring 2019 cycle of work, the two pairs of competing measures were resubmitted for maintenance of endorsement. Because the measures are competing (defined as "same concepts for measure focus—target process, condition, event, outcome and targeting the same population"), the Committee should attempt to pick one measure from each pair to recommend for endorsement. This is termed as "best-in-class," and NQF has developed an algorithm to assist in selection.

NQF's Guidance on Competing Measures

The following tables and figure are from NQF's 2018 Measure Evaluation Criteria Guidance.

Table 8. Related Versus Competing Measures

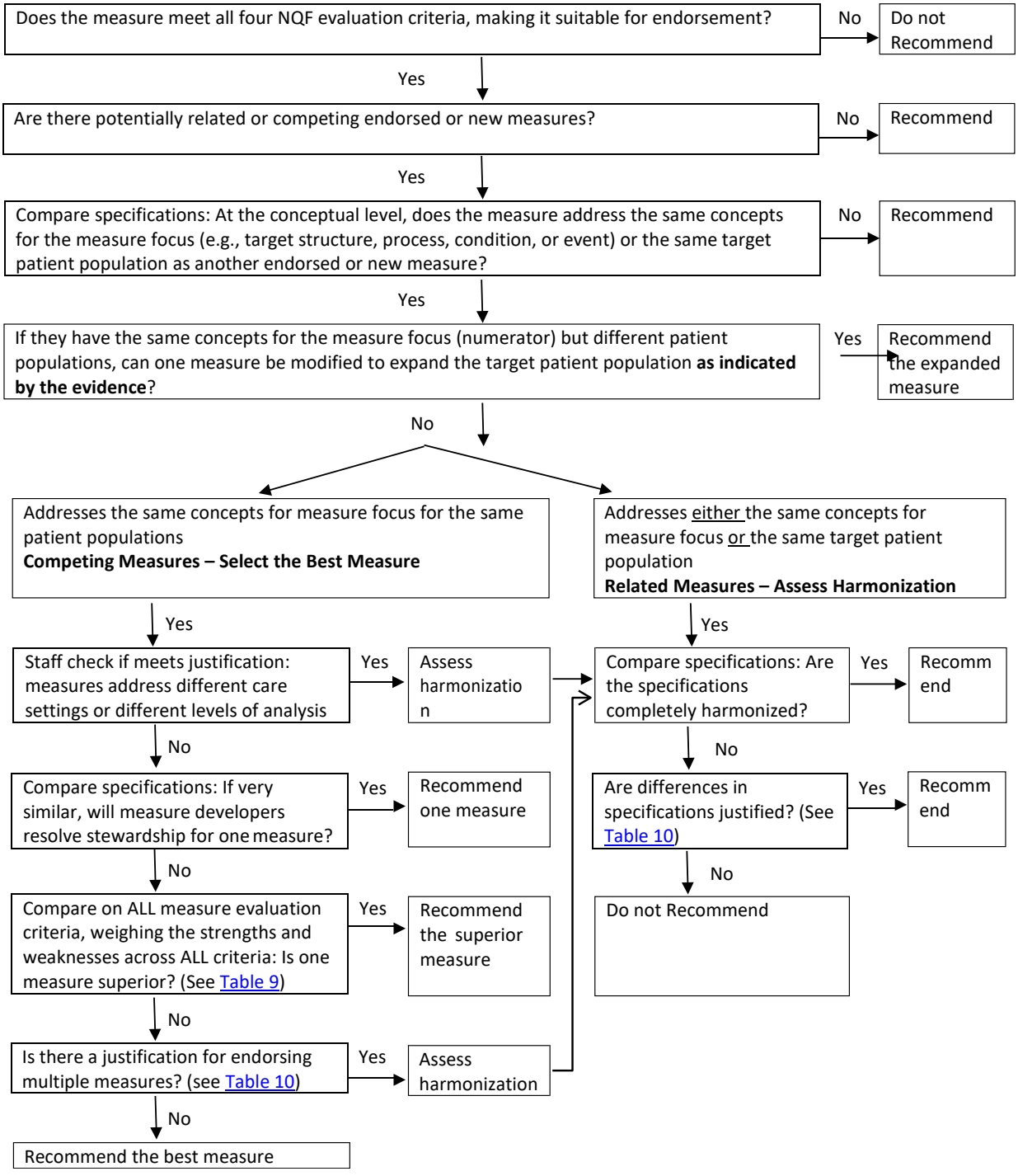
	Same concepts for measure focus—target process, condition, event, outcome	**Different concepts for measure focus—target process, condition, event, outcome**
Same target patient population	**Competing measures—Select best measure** from competing measures or justify endorsement of additional	**Related measures—Harmonize** on target patient population or justify differences.
Different target patient population	**Related measures—Combine** into one measure with expanded target patient population or justify why different harmonized measures are needed.	Neither harmonization nor competing measure issue

Table 9. Evaluating Competing Measures for Superiority or Justification for Multiple Measures

Steps	**Evaluate Competing Measures**
1. Determine if need to compare measures for superiority	Work through the steps in the algorithm (Figure 1) to determine if need to evaluate competing measures for superiority (i.e., two or more measures address the same concepts for measure focus for the same patient populations)
2. Assess Competing Measures for Superiority by weighing the strengths and weaknesses across ALL NQF evaluation criteria	**Because the competing measures have already been determined to have met NQF's criteria for endorsement, the assessment of competing measures must include weighing the strengths and weaknesses across ALL the criteria and involves more than just comparing ratings.** (For example, a decision is not based on just the differences in scientific acceptability of measure properties without weighing the evaluation of importance to measure and report, usability, and feasibility as well.) **Evidence, Performance Gap—Importance to Measure and Report:** Competing measures generally will be the same in terms of the evidence for the focus of measurement (1a). However, due to differences in measure construction, they could differ on performance gap (1b). Compare measures on opportunity for improvement (1b) **Reliability and Validity—Scientific Acceptability of Measure Properties:** Compare evidence of reliability (2a1-2a2) Compare evidence of validity, including threats to validity (2b1-2b6) Untested measures cannot be considered superior to tested measures because there would be no empirical evidence on which to compare reliability and validity. However, a new measure, when tested, could ultimately demonstrate superiority over an endorsed measure and the NQF endorsement maintenance cycles allow for regular submission of new measures. Compare and identify differences in specifications **All else being equal on the criteria and subcriteria**, the preference is for: Measures specified for the broadest application (target patient population as indicated by the evidence, settings, level of analysis) Measures that address disparities in care when appropriate

Steps	Evaluate Competing Measures
	Feasibility: Compare the ease of data collection/availability of required data **All else being equal on the criteria and subcriteria**, the preference is for: Measures based on data from electronic sources Clinical data from EHRs Measures that are freely available **Usability and Use:** Compare evidence of the extent to which potential audiences (e.g., consumers, purchasers, providers, policymakers) are using or could use performance results for both accountability and performance improvement. **All else being equal on the criteria and subcriteria**, the preference is for: Measures used in at least one accountability application Measures with the widest use (e.g., settings, numbers of entities reporting performance results) Measures for which there is evidence of progress towards achieving high-quality efficient healthcare for individuals or populations The benefits of the measure outweigh any unintended negative consequences to individuals or populations **After weighing the strengths and weaknesses across ALL criteria, identify if one measure is clearly superior and provide the rationale based on the NQF criteria.**

Figure 1. Addressing Competing Measures and Harmonization of Related Measures in the NQF Evaluation Process



2015 Project: History and Conditions

Related and Competing Measures

NQF requires that committees consider whether measures are related (either the same measure focus or the same target population) or competing (both the same measure focus and the same target population) with other measures in the portfolio. NQF staff identified seven sets of measures as related and two sets of measures as competing during their preliminary analysis developed by UDSMR and CMS. Comments were made by proponents of the UDSMR measures (based on the FIM® tool) and by proponents of the CMS measures (based on the Continuity Assessment Record and Evaluation [CARE] tool). Following the Committee's final recommendations on the consensus not reached and not recommended measures, the Committee convened via web meeting on May 1, 2015, to discuss the related and competing measures. The Committee agreed that the seven sets of measures identified by NQF are related but did not make recommendations for harmonization. In their discussions, the Committee indicated that the related measures either addressed different populations or were varied enough in their focus area to support moving the measures forward through the endorsement process. The Committee members considered two pairs of UDSMR and CMS measures as competing and as such were asked to complete a voting survey after the call. The competing measures included:

- **2633** Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients (CMS) and **2286** Functional Change: Change in Self Care Score (UDSMR); and
- **2634** Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients (CMS) and **2321** Functional Change: Change in Mobility Score (UDSMR).

The Committee agreed that each pair of measures was competing, but could not come to consensus on "best-in-class" in either case. Therefore, both pairs of measures were recommended for endorsement to the Consensus Standards Approval Committee (CSAC) as competing but with consensus not reached on "best-in-class." Committee members provided the following rationale for not choosing a "best-in-class" in either set.

- Measures 2286 and 2321 (UDSMR) have a long history of utilization nationally, and are specified for all adult patients, not just the Medicare population. Significant costs (personnel re-training, software systems for capturing data) would accompany a switch to another measure, without clear added benefit to the institutions involved in rehabilitation.
- One measure in each pair is "tried and true," and the other is emerging with a good possibility of becoming superior over time.
- One measure in each pair is based on the FIM® and has a long history; healthcare staff across the country are trained and familiar with it; and it would be a major upheaval not to endorse this measure. The other measure in each pair is based on the CARE tool and was developed using more contemporary science, is designed to

cut across settings of post-acute care and has had significant investment by CMS in its development and refinement.

- It is hard to say whether one is superior at this time. By allowing both measures to continue for the time being, CMS and other payers will be able to employ both measures and continue to experience how they work in practice, perhaps building an evidence base for future selection of one superior measure.

2015 Process

These four measures were recommended for endorsement by the Standing Committee after considerable public comment, member voting, and additional information provided by measure developers. The Standing Committee was unable to select a best-in-class for either pair of competing measures (2633 versus 2286 and 2634 versus 2321). The two UDSMR measures (2286 and 2321) were recommended for endorsement with 71 percent of NQF member councils approving. The councils were unable to reach consensus for the two CMS measures (2633 and 2634) with only 56 percent of councils approving the measures.

In their initial vote in June, the CSAC voted to recommend the two UDSMR measures, while the two CARE tool measures only received 56 percent approval (below the required 60 percent threshold for CSAC approval). Based on the rationale provided by CSAC members, the CMS IRF measures were not approved largely due to competing measure concerns.

In their reconsideration vote in September 2015, 12 out of 13 CSAC members, or 92 percent, voted to approve endorsement for the four measures with conditions for specific updates.

Update Requirements:

UDSMR	CMS
• Provide information about how the inclusion or exclusion of cognitive items impacts the overall assessment of the patient. • Provide updated measure-level testing for reliability and validity given that all the measures are new. There is particular interest in measure performance/scientific acceptability across care settings beyond IRF. • Provide information about costs associated with use of the FIM Instrument, respective software and tools, and costs of ongoing training in order to accurately use the FIM Instrument.	• Provide information about how the inclusion or exclusion of cognitive items impacts the overall assessment of the patient. • Provide updated measure-level testing for reliability and validity given that all the measures are new and will be implemented in 2016. • Provide data on comparison of the competing measure results to gain an understanding of which scale is more reliable, valid, and feasible. • Provide a summary of qualitative data gathered during rulemaking process including perceived benefits from the field for instruments that cut across settings.

Action Items for Committee on June 20

Evaluate each of the four measures against NQF's endorsement criteria; if a measure passes all of the must-pass criteria, the Committee will vote on a recommendation for endorsement.

- If both measures in a pair pass all of the must-pass criterion and are recommended for endorsement, the Committee will then move to the best-in-class discussion.
- If one of the measures does not pass all of the must-pass criterion or is not recommended by the Committee on the overall vote and the other does, the passing measure is automatically considered the best in class.
- If neither measure passes a must-pass or the recommendation for endorsement, there is no competing measures discussion.
- If consensus is not reached (CNR) on one of the measures, but the other measure passes, the best-in-class discussion is tabled until after the Committee's second discussion and vote on the CNR measure.
- If consensus is not reached on either of the competing measures discussion, the best-in-class discussion and vote is tabled until the post-comment call, assuming the Committee is able to come to consensus on both measures at that time.

If consensus on the competing decision is not reached at the in-person meeting, it will be discussed on the post-comment call.

Comments and Their Disposition

During the pre-meeting commenting period, NQF received two comments from one individual who was not an NQF member pertaining to the competing measures. The comments were submitted on 2286 and 2231 and supported the endorsement of the CARE tool-based measures over the FIM tool-based measures.

Comments submitted during this comment period are included in Appendix C.

Question for Committee to Consider:

2286 Change in Self Care Function and Measure 2321 Change in Mobility Function:

- These measures will no longer be in use in IRF-PAI as of October 2019. NQF's maintenance criteria require that maintenance measures be in use for continued endorsement. The developer's submission does not include information about the future uses of these measures. Does the Committee have any concerns about the current or future use of these measures?

Appendix A: Additional Information Submitted by Developers

UDSMR: Measures 2286 and 2321

Additional information requested on May 30, 2019 by NQF Patient Experience and Function Committee for Measure #2286 Change in Self Care Function and Measure # 2321 Change in Mobility Function:

- A year over year assessment of performance on the measure, preferably for the last three years and longer if possible
- Any attempts made at harmonization of the measures
- Approaches that could be implemented to harmonize

For Measure #2286 Change in Self Care Function:

- A year over year assessment of performance on the measure, preferably for the last three years and longer if possible

Self Care Measure: Distribution of Facilities (Number and Percentages) by Mean Change in Self Care Score by Quartile by Year

Total Number of Facilities by Year	Below 25th%	50th%	75th% and Above
2015 (n=840)	112 (13.3%)	490 (58%)	238 (28.3%)
2016 (n=838)	86 (10%)	479 (57%)	273 (32.6%)
2017 (n=836)	37 (4%)	489 (58%)	310 (37.1%)
2018 (n=914)	77 (8%)	477 (52%)	360 (39.4%)

The table above displays facility-level improvement over time, as captured by the change in self care measure. The percentage of facilities in the lowest quartile, below 25th%, has decreased from 13.3% to 8% while the percentage of facilities in the top quartile has increased from 28.3% to 39.4% from 2015 to 2018.

Mean Change in Self Care Score by Quartile by Year

2015				2016				2017				2018			
Quartile	**Mean**	**N**	**Std. Dev.**	**Quartile**	**Mean**	**N**	**Std. Dev.**	**Quartile**	**Mean**	**N**	**Std. Dev.**	**Quartile**	**Mean**	**N**	**Std. Dev.**
Below 25%	3.87	124292	4.03	Below 25%	3.80	120060	4.09	Below 25%	4.58	142637	4.24	Below 25%	3.75	124485	4.13
50th%	11.07	120297	1.40	50th%	11.09	120886	1.40	50th%	11.54	103348	1.11	50th%	11.12	132458	1.40
75th%	15.88	110277	1.41	75th%	15.89	115091	1.40	75th%	15.89	119853	1.40	75th%	15.89	134310	1.40
Above 75th%	23.09	106062	3.83	Above 75th%	23.27	117189	3.96	Above 75th%	23.33	123104	4.02	Above 75th%	23.28	137415	3.98
Total	13.04	460928	7.59	Total	13.42	473226	7.71	Total	13.54	488942	7.74	Total	13.76	528668	7.69
Sig. =	p<.001	Eta2 =	0.85	Sig. =	p<.001	Eta2 =	0.85	Sig. =	p<.001	Eta2 =	0.83	Sig. =	p<.001	Eta2 =	0.85

The table above displays the mean change in self care score and standard deviation by quartile by year. Results of an ANOVA were statistically significant, whereby there were significant differences in mean change scores by quartile. These scores were significantly different for each year and they were relatively stable and consistent over time, for instance, the mean score at the 50th % was 11.1 in 2015, 2016 and 2018 and it was 11.5 in 2017 (only a very slight change from the other years), this pattern holds for the other quartiles. Furthermore, the eta^2 value is very strong for each year, lending supporting evidence that the differences in mean scores by quartile are truly different and the measure account for a large proportion of variance between groups (between 83% (2017) to 85%).

For Measure #2321 Change in Mobility Function:

- A year over year assessment of performance on the measure, preferably for the last three years and longer if possible

Mobility Measure: Distribution of Facilities (Number and Percentages) by Mean Change in Mobility Score by Quartile by Year

Total Number of Facilities by Year	Below 25th%	50th%	75th% and Above
2015 (n=840)	21 (2.5%)	679 (81%)	140 (16.7%)
2016 (n=838)	14 (1.7%)	656 (78.3%)	168 (20%)
2017 (n=836)	10 (1.2%)	624 (74.6%)	202 (24.2%)
2018 (n=914)	13 (1.4%)	680 (74.4%)	221 (24.2%)

The table above displays facility-level improvement over time, as captured by the change in mobility measure. The percentage of facilities in the lowest quartile, below 25th%, has decreased from 2.5% to 1.4% while the percentage of facilities in the top quartile has increased from 16.7% to 24.2% from
2015 to 2018.

Mean Change in Mobility Score by Quartile by Year

2015				2016				2017				2018			
Quartile	**Mean**	**N**	**Std. Dev.**	**Quartile**	**Mean**	**N**	**Std. Dev.**	**Quartile**	**Mean**	**N**	**Std. Dev.**	**Quartile**	**Mean**	**N**	**Std. Dev.**
Below 25%	2.27	115051	2.18	Below 25%	2.21	115306	2.21	Below 25%	2.79	135779	2.59	Below 25%	2.17	125063	2.22
50th%	7.61	122390	1.11	50th%	7.62	121357	1.11	50th%	8.64	140216	1.11	50th%	7.63	131350	1.11
75th%	11.44	135447	1.11	75th%	11.46	140117	1.11	75th%	11.49	76065	0.50	75th%	11.46	157635	1.11
Above 75th%	15.81	88040	1.77	Above 75th%	15.84	96446	1.80	Above 75th%	15.20	136882	2.03	Above 75th%	15.89	114620	1.82
Total	8.97	460928	4.99	Total	9.11	473226	5.10	Total	9.30	488942	5.08	Total	9.27	528668	5.11
Sig. =	p<.001	Eta2 =	0.9	Sig. =	p<.001	Eta2 =	0.91	Sig. =	p<.001	Eta2 =	0.87	Sig. =	p<.001	Eta2 =	0.9

The table above displays the mean change in mobility score and standard deviation by quartile by year. Results of an ANOVA were statistically significant, whereby there were significant differences in mean change scores by quartile. These scores were significantly different for each year and they were relatively stable and consistent over time, for instance, the mean score at the 75th % was 11.4 for each of the four years, this pattern holds for the other quartiles as well. Furthermore, the eta^2 value is very strong for each year, lending supporting evidence that the differences in mean scores by quartile are truly different and the measure accounts for a large proportion of variance between groups (between 87% (2017) to 91%).

For Measure #2286 Change in Self Care Function and Measure # 2321 Change in Mobility Function:

- Any attempts made at harmonization of the measures
- Approaches that could be implemented to harmonize

Measure #2286 is similar to CMS Measure #2633, except Measure #2286 is intended for all patients ages 18 and older receiving post acute care (at an IRF, SNF or LTAC facility) and Measure #2233 is intended for use among Medicare patients only (of which the majority is persons aged 65 and older) receiving care at an IRF. Furthermore, Measure #2286 includes both physical and cognitive functional items whereas Measure #2633 includes only physical items and does not include any items specific to cognitive self-care function. Measure 2286 includes 8 self-care items (6 physical and 2 cognitive) rated on a 7 level scale, clinicians rate the patient´s lowest actual observed score over the past 24 hour period (if patient is independent with toileting while awake but needs assistance in the middle of the night the rating would be the lowest/middle of the night score for the item), all items are to be rated at admission and at discharge, there are no codes for missing/do not apply. Measure 2633 includes 7 self-care items all of which are physical items, and uses a 6 level rating scale (1-6) which includes options for not assessing each item, thus allows for missing responses (ex. not applicable/ patient refused/ not attempted due to safety). For measure 2633, the patient´s usual performance is used as the basis of the score whereby if a patient were independent in toileting during the day but needed assistance in the middle of the night the score would be independent as there would be more frequent independent episodes throughout the day opposed to a single instance during the night. The two measures use different rating scales and different assessment rules and when trying to determine a patient´s actual level of function, a 6 level scale is less sensitive than a 7 level scale as there is less ´room´ to demonstrate change over time captured in the 6 level rating scale. Additionally, if clinicians are using the patient's functional level to determine patient discharge setting, using patient 'usual performance´ may portray a higher level of function than truly exists for the patient, whereby if it is believed the patient is independent in certain items but does in fact need assistance at certain times of day or in some instances, and there are not provisions in place to provide the care, the patient is at risk for a fall or readmission to inpatient care if a caregiver or attendant is not with the patient to provide the assistance (such as in the example of

toileting used previously). Furthermore, the inclusion of multiple 'missing' data options for each item to be used at both admission and at discharge lends the possibility for data that is not able to be interpreted, if an item is not rated at admission because the patient refused but is rated at discharge, of what value is this information? It is unknown if the patient would have been rated the exact same at admission, thus no change actually occurred from admission to discharge, of if there were an improvement, it would not be captured, or if there was a decline in function, this too is unknown, so if an item is not applicable (or not safe for administration at admission) than it lends question as to why it is included in the measure at all and if it is applicable, allowing missing values adds to the clinical data collection burden without any benefit to the patient as any other values collected cannot be interpreted directly when an item was missing at another point in time (an admission rating but no discharge rating or vice versa). Predictive models at the measure level require complete data so even if one value is missing for one item the entire case is dropped from the analytical model so the facility level outcome data would be impacted by the missing values as well as the patient level outcomes data. For the above detailed reasons, Measures #2286 and #2633 are conceptually and fundamentally different and cannot be harmonized. Measure #2286 is best in class as it has a long-standing history of use in addition to evidence of high reliability, validity and predictive ability.

Measure #2321 is similar to CMS Measure #2634, however Measure #2634 is only intended for Medicare patients (majority of which are age 65 or older) treated at an IRF whereas Measure #2321 is intended for all patients age 18 and older receiving post acute care at an IRF, SNF or LTAC facility. Measure #2321 includes four mobility items, whereby Measure #2634 includes 15 mobility items, several of which are very similar in nature and may add little to no value in assessing patient function (the individual contribution of each item within the measure could be assessed using rasch analysis and/or factor analysis). There are infinite number of items that could be assessed but each of these add to patient and clinician assessment burden so it is paramount that each item is truly capturing the construct of interest (mobility function) and that the measure as a whole is parsimonious. Furthermore, several of the items in measure #2634 are not feasible for patients in an inpatient setting, such as: car transfer, walk on uneven surfaces, bend to pick up an object while standing, as these items are not safe for the large majority of patients admitted to an inpatient rehabilitation facility upon admission as most cannot perform these tasks and among those who could, most clinicians would agree that it would be unsafe to even attempt. This point is acknowledged considering there are four missing codes for all of the CMS mobility items in measure #2634 (patient refused, not applicable, not attempted due to safety concerns and not attempted due to environmental limitations). Measure #2321, however, is applicable for all adult patients and is intended to be assessed on all patients at both admission and discharge. If an item is not applicable at admission, a change score cannot be computed and true assessment of patient and facility outcomes may be biased based on the missing data. Furthermore, true validation of the measure requires complete data for all items within the measure, otherwise cases with even just one item missing are eliminated from the statistical model. This may result in a large amount of missing data compared to the total number of cases assessed and the

results of the analysis would be biased to include only complete cases with no missing data, these cases are likely very different (and much higher functioning, if a patient can walk on an uneven surface at admission to an IRF) than other patients where the given item(s) was not attempted at admission (more typical in terms of the type of patient admitted to an IRF). Measure #2321 and #2634 use different rating scales and different assessment rules, a 6 level scale is used for measure #and patient 'usual performance´ is the basis of the score whereas a 7 level scale is used for measure #2321 and actual, observed performance is the basis of the score. The use of 'usual performance' in assessment of function may portray a higher level of function than truly exists for the patient, whereby if it is believed the patient is independent in certain items but does in fact need assistance at certain times of day or in some instances, and there are not provisions in place to provide the care, the patient is at risk for a fall or readmission to inpatient care if a caregiver or attendant is not with the patient to provide the assistance. For the above detailed reasons, measures #2321 and #2634 are conceptually and fundamentally different and cannot be harmonized. Measure #2321 is best in class as it has a long-standing history of use in addition to evidence of high reliability, validity and predictive ability.

CMS & RTI Memo 1

To: **National Quality Forum Patient Experience and Function Committee**

From: RTI International and the Centers for Medicare and Medicaid Services

Date: 06/06/2019

Subject: Quality Measure Scores for Fiscal Year 2017 and Fiscal Year 2018 for IRF Functional Outcome Measures (NQF #2633, NQF #2634, NQF #2635, NQF #2636)

Data collection for the functional outcome measures began October 1, 2016 and providers have 4.5 months to review and correct their data submitted to Centers for Medicare and Medicaid Services (CMS) before the data are finalized. Using the national IRF-PAI data, we provide fiscal year 2017 and fiscal year 2018 quality measure scores using 12 months of data, as well as quality measure scores by quarter. The fiscal year 2017 IRF-PAI data includes data for Medicare patients discharged from IRFs between October 1, 2016 – September 30, 2017 (N=490,032) whereas the fiscal year 2018 IRF-PAI data includes data for Medicare patients discharged between October 1, 2017 – September 30, 2018 (N=504,927) before exclusion criteria are applied.

Quality measure score distributions by 12-month time period:

1. Fiscal year 2017 (October 1, 2016 – September 30, 2017) (n=1,119 providers)
2. Fiscal year 2018 (October 1, 2017 – September 30, 2018) (n=1,111 providers)

Quality measure score distributions by quarter (October 1, 2016 – September 30, 2018; 8 quarters):

1. Quarter 4, 2016 (n=1,103)
2. Quarter 1, 2017 (n=1,105)
3. Quarter 2, 2017 (n=1,107)
4. Quarter 3, 2017 (n=1,107)
5. Quarter 4, 2017 (n=1,096)
6. Quarter 1, 2018 (n=1,096)
7. Quarter 2, 2018 (n=1,101)
8. Quarter 3, 2018 (n=1,093)

Overall, between fiscal year 2017 and fiscal year 2018, mean scores remained stable or marginally fluctuated for all four functional outcome quality measures (NQF #2633, NQF #2634, NQF #2635, NQF #2636). Analysis of data over the eight quarters found slight increases in mean scores for all four quality measures.

Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients (NQF #2633)

The mean quality measure score (12-months) increased slightly between fiscal year 2017 (mean: 11.4 units of change in self-care; standard deviation: 1.7) and fiscal year 2018 (mean: 11.6 units of change in self-care; standard deviation: 1.8) (Table 1). Quality measure scores by decile show slight increases in quality measure scores for each decile between fiscal year 2017 and fiscal year 2018. When we analyzed data by quarter, we observed mean scores increased from 11.3 units of change in self-care to 11.7 units of change in self-care across eight quarters (Q4, 2016 – Q3, 2018) (Table 2).

Table 1. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients (NQF #2633): Quality Measure Scores for Fiscal Year 2017 and Fiscal Year 2018

	Fiscal Year 2017 (12 months)	Fiscal Year 2018 (12 months)
Number of Facilities	1,119	1,111
Mean score	11.4	11.6
Standard deviation	1.7	1.8
Interquartile range	2.2	2.3
Score deciles		
1st decile	8.3	8.5
2nd decile	9.7	9.9
3rd decile	10.3	10.5
4th decile	10.7	11.0
5th decile	11.2	11.4
6th decile	11.5	11.9
7th decile	12.0	12.3
8th decile	12.5	12.8
9th decile	13.1	13.4
10th decile	14.3	14.7
Minimum	5.1	4.5
Maximum	17.0	17.4

Note: Scores are reported as units of change in self-care; Providers with < 20 stays during the 12-month testing period are excluded.

Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV64, MV74).

Table 2. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients (NQF #2633): Quality Measure Score Distributions by Quarter

	Number of Facilities	Mean Score	Standard Deviation	Interquartile Range	Minimum	Maximum
Fiscal Year 2017						
Oct-Dec 2016	1,103	11.3	1.9	2.5	2.8	18.9
Jan-Mar 2017	1,105	11.4	1.9	2.3	3.8	19.2
Apr-Jun 2017	1,107	11.5	1.9	2.4	4.4	17.8
Jul-Sept 2017	1,107	11.5	1.9	2.6	3.3	18.5
Fiscal Year 2018						
Oct-Dec 2017	1,096	11.6	1.9	2.3	1.4	18.3
Jan-Mar 2018	1,096	11.6	1.9	2.5	4.7	18.9
Apr-Jun 2018	1,101	11.7	2.0	2.6	1.8	18.4
Jul-Sept 2018	1,093	11.7	2.0	2.5	3.6	18.2

Note: Scores are reported as units of change in self-care; Providers with < 20 stays during the 12-month testing period are excluded. Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV64, MV74).

Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients (NQF #2634)

The mean quality measure score (12-months) remain the same between fiscal year 2017 (mean: 28.2 units of change in mobility; standard deviation: 4.6) and fiscal year 2018 (mean: 28.2 units of change in mobility; standard deviation: 4.7) (Table 1). Quality measure scores by decile show slight increases in quality measure scores for each decile between fiscal year 2017 and fiscal year 2018. When we analyzed data by quarter, we observed mean scores increased slightly from 27.9 units of change in mobility to 28.3 units of change in mobility across eight quarters (Q4, 2016 – Q3, 2018) (Table 2).

Table 3. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients (NQF #2634): Quality Measure Scores for Fiscal Year 2017 and Fiscal Year 2018

	Fiscal Year 2017 (12 months)	Fiscal Year 2018 (12 months)
Number of Facilities	1,119	1,111
Mean score	28.2	28.2
Standard deviation	4.6	4.7
Interquartile range	6.0	6.4

	Fiscal Year 2017 (12 months)	Fiscal Year 2018 (12 months)
Score deciles		
1st decile	20.4	20.3
2nd decile	23.5	23.5
3rd decile	25.1	25.0
4th decile	26.3	26.3
5th decile	27.5	27.5
6th decile	28.5	28.7
7th decile	29.6	30.0
8th decile	31.2	31.4
9th decile	33.0	33.2
10th decile	36.4	36.6
Minimum	13.7	13.4
Maximum	52.6	51.1

Note: Scores are reported as units of change in mobility; Providers with < 20 stays during the 12-month testing period are excluded.

Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV64, MV74).

Table 4. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients (NQF #2634): Quality Measure Score Distributions by Quarter

	Number of Facilities	Mean Score	Standard Deviation	Interquartile Range	Minimum	Maximum
Fiscal Year 2017						
Oct-Dec 2016	1,103	27.9	5.1	6.6	8.4	55.8
Jan-Mar 2017	1,105	28.1	4.9	6.5	13.3	55.8
Apr-Jun 2017	1,107	28.4	5.1	6.6	9.1	52.2
Jul-Sept 2017	1,107	28.4	5.1	6.8	12.7	47.1
Fiscal Year 2018						
Oct-Dec 2017	1,096	28.0	5.1	7.1	3.3	52.2
Jan-Mar 2018	1,096	28.2	5.2	6.7	3.5	48.6
Apr-Jun 2018	1,101	28.3	5.2	7.2	10.5	54.4
Jul-Sept 2018	1,093	28.3	5.0	6.7	11.8	54.0

Note: Scores are reported as units of change in mobility; Providers with < 20 stays during the 12-month testing period are excluded.

Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV64, MV74).

Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients (NQF #2635)

The mean quality measure score (12-months) remained the same between fiscal year 2017 (mean: 55.9 percent; standard deviation: 14.2) and fiscal year 2018 (mean: 55.9 percent; standard deviation: 14.5) (Table 1). Quality measure scores by decile show slight increases in quality measure scores for each decile between fiscal year 2017 and fiscal year 2018. When we analyzed data by quarter, we observed mean scores increased slightly from 55.2 percent to 56.4 percent across eight quarters (Q4, 2016 – Q3, 2018) (Table 2).

Table 5. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients (NQF #2635): Quality Measure Scores for Fiscal Year 2017 and Fiscal Year 2018

	Fiscal Year 2017 (12 months)	Fiscal Year 2018 (12 months)
Number of Facilities	1,119	1,111
Mean score	55.9	55.9
Standard deviation	14.2	14.5
Interquartile range	19.0	20.1
Score deciles		
1st decile	29.0	29.0
2nd decile	41.1	40.7
3rd decile	46.9	46.3
4th decile	51.0	50.4
5th decile	54.8	54.2
6th decile	58.2	58.7
7th decile	61.8	62.7
8th decile	66.2	66.6
9th decile	70.7	71.0
10th decile	78.8	79.2
Minimum	6.7	10.1
Maximum	91.2	94.5

Note: Scores are reported as percentiles; Providers with < 20 stays during the 12-month testing period are excluded.

Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV65, MV75).

Table 6. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients (NQF #2635): Quality Measure Score Distributions by Quarter

	Number of Facilities	Mean Score	Standard Deviation	Interquartile Range	Minimum	Maximum
Fiscal Year 2017						
Oct-Dec 2016	1,103	55.2	15.8	22.6	0.0	100.0
Jan-Mar 2017	1,105	55.5	15.6	21.1	7.1	100.0
Apr-Jun 2017	1,107	56.1	15.9	22.5	0.0	100.0
Jul-Sept 2017	1,107	55.8	16.1	20.8	0.0	100.0
Fiscal Year 2018						
Oct-Dec 2017	1,096	55.1	15.9	21.1	5.3	100.0
Jan-Mar 2018	1,096	55.6	16.1	21.5	0.0	100.0
Apr-Jun 2018	1,101	56.1	16.4	23.5	0.0	100.0
Jul-Sept 2018	1,093	56.4	16.1	23.0	5.9	93.2

Note: Scores are reported as percentiles; Providers with < 20 stays during the 12-month testing period are excluded.

Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV65, MV75).

Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Mobility Score for Medical Rehabilitation Patients (NQF #2636)

The mean quality measure score (12-months) decreased between fiscal year 2017 (mean: 50.8 percent; standard deviation: 14.7) and fiscal year 2018 (mean: 50.1 percent; standard deviation: 14.8) (Table 1). Quality measure scores by decile show slight increases in quality measure scores for each decile between fiscal year 2017 and fiscal year 2018. When we analyzed data by quarter, we observed mean scores varied across the eight quarters (Q4, 2016 – Q3, 2018) (Table 2) with an overall slight increase from 50.0 percent to 50.4 percent.

Table 7. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Mobility Score for Medical Rehabilitation Patients (NQF #2636): Quality Measure Scores for Fiscal Year 2017 and Fiscal Year 2018

	Fiscal Year 2017 (12 months)	Fiscal Year 2018 (12 months)
Number of Facilities	1,119	1,111
Mean score	50.8	50.1
Standard deviation	14.7	14.8
Interquartile range	21.0	21.7
Score deciles		
1st decile	24.5	29.0
2nd decile	34.8	40.7
3rd decile	40.4	46.3
4th decile	45.2	50.4
5th decile	49.3	54.2
6th decile	53.0	58.7
7th decile	57.1	62.7
8th decile	61.5	66.6
9th decile	66.7	71.0
10th decile	75.2	79.2
Minimum	8.1	9.8
Maximum	94.4	92.2

Note: Scores are reported as percentiles; Providers with < 20 stays during the 12-month testing period are excluded.

Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV65, MV75).

Table 8. Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Mobility Score for Medical Rehabilitation Patients (NQF #2636): Quality Measure Score Distributions by Quarter

	Number of Facilities	Mean Score	Standard Deviation	Interquartile Range	Minimum	Maximum
Fiscal Year 2017						
Oct-Dec 2016	1,103	50.0	16.4	22.6	0.0	95.7
Jan-Mar 2017	1,105	50.3	16.2	22.8	8.8	100.0
Apr-Jun 2017	1,107	50.8	16.6	23.9	0.0	100.0
Jul-Sept 2017	1,107	50.9	16.5	23.8	0.0	93.9
Fiscal Year 2018						
Oct-Dec 2017	1,096	49.5	16.2	23.8	0.0	96.1
Jan-Mar 2018	1,096	49.9	16.6	23.5	0.0	100.0
Apr-Jun 2018	1,101	50.1	16.8	23.7	0.0	100.0
Jul-Sept 2018	1,093	50.4	16.2	23.4	0.0	98.1

Note: Scores are reported as percentiles; Providers with < 20 stays during the 12-month testing period are excluded.

Source: RTI analysis of IRF-PAI October 2016 – September 2018 (Program reference: MV65, MV75).

CMS & RTI Memo 2

To: **National Quality Forum Patient Experience and Function Committee**

From: RTI International and the Centers for Medicare and Medicaid Services

Date: 06/06/2019

Subject: Harmonization of Self-Care and Mobility Quality Measures

The purpose of this memo is to describe features of the Change in Self-Care (NQF #2633) and Change in Mobility (NQF #2634) quality measures and provide notes about how these quality measures are harmonized with related and competing measures. The information in this memo is included in the Measure Information Forms for NQF #2633 and NQF #2634 (submitted to NQF as part of measure endorsement maintenance on 4/22/2019). We have excerpted key content from the Measure Information Forms and presented some information in tables for this memo. We would be happy to address any questions from the Committee about harmonization issues.

I. INPATIENT REHABILITATION FACILITY (IRF) FUNCTIONAL OUTCOME MEASURE: CHANGE IN SELF-CARE SCORE FOR MEDICAL REHABILITATION PATIENTS (NQF #2633)

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0174: Improvement in bathing
0175: Improvement in bed transferring
0426: Functional status change for patients with Shoulder impairments
0427: Functional status change for patients with elbow, wrist and hand impairments
0428: Functional status change for patients with General orthopaedic impairments
0688: Percent of Residents Whose Need for Help with Activities of Daily Living Has Increased (long stay)
2286: Functional Change: Change in Self Care Score
2287: Functional Change: Change in Motor Score
2613: CARE: Improvement in Self Care
2643: Average change in functional status following lumbar spine fusion surgery
2769: Functional Change: Change in Self Care Score for Skilled Nursing Facilities
2775: Functional Change: Change in Motor Score for Skilled Nursing Facilities
2776: Functional Change: Change in Motor Score in Long Term Acute Care Facilities
2777: Functional Change: Change in Self Care Score for Long Term Acute Care Facilities

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

All the listed measures address the same topic, function, but the target populations for most of these measures is not the IRF patient population. For example, measures are used for patients/residents treated in outpatient settings, home care, skilled nursing facilities, long-stay nursing homes, and long-term care hospitals. One measure has been previously identified by NQF staff as a competing measure: Functional Change: Change in Self Care Score (NQF #2286).

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

The NQF and the Patient Experience and Function Standing Committee may choose to endorse both competing measures, because both provide value. If NQF and the committee believe that only one measure should be endorsed as "best-in-class," we offer a list of the strengths of our measure below as well as a comparison of feasibility, usability and use for consideration. Specifically, we describe the similarities and important differences between this change in self-care measure and the listed related and competing measures (See 5.1.a). We note that several features of this measure (e.g., the data elements, many of the risk adjustors, and the risk-adjustment approach) are the same as or aligned with the specifications of several of the other endorsed measures. Therefore, we believe that the specifications for this measure incorporate the best features of all endorsed related and competing measures, and, as a whole, represents the "best in class" for measuring change in self-care for IRFs.

This Change in Self-Care (NQF #2633) measure was developed by building on the most recent science related to measurement of patient functioning and quality measure development. The latest science and scholarly literature, clinical thinking, and expert input on functional assessment and quality measurement was combined with a cross-setting design and purpose in mind. Specifications were discussed with stakeholders and experts, pilot tested, and analyzed throughout the development process, as described in the Testing form.

Table 1. Features of the Measure and Notes about Harmonization: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients (NQF #2633)

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
Functional Assessment Data Elements	
1. Cross-Setting Design The functional assessment data elements for this measure, included in Section GG: Functional Abilities and Goals, were designed and tested with a cross-setting purpose in mind to ensure that data may be collected by clinicians in various post-acute and acute care settings. This enhances the cross-setting validity and reliability of quality measures that use these data. Standardization of self-care and mobility data elements across post-acute care settings has been an important goal for policymakers and included in the IMPACT Act of 2014.	We note that another measure focused on improvement in self-care, Related Measure NQF #2613, also use the data elements from Section GG: Functional Abilities and Goals as part of their performance measure with the rationale that the data elements were developed for cross-setting use and that the data elements are standardized.

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
2. Clinician Observation To determine a patient's functional ability, providers are instructed to code the data elements in Section GG: Functional Abilities and Goals primarily based on clinical observation. Specifically, a qualified clinician will assess the patient's performance based on direct observation, as well as gather input from reports from other clinicians, care staff, or family as well as the patient's self-report. Typically, an interdisciplinary team of qualified clinicians is involved in assessing the patient and CMS provides guidance through manuals, training programs, and help desk responses to support providers in collecting accurate functional assessment data.	We note that the Competing Measure NQF #2286 and Related Measures NQF #2613, #2769, #2775, #2776, and #2777 also use clinician observation to assess and code a patient's functional abilities.
3. Functional Assessment Data Elements Capture A Range of Functioning The functional assessment data elements and associated rating scale were designed to build on the existing science of functional assessment, which included a review of the strengths and limitations of existing instruments. The inclusion of 7 self-care data elements allows for the measurement of a wide range of patient functioning and thus the opportunity to demonstrate gains in a variety of functional activities. Patients may be expected to make varying amounts of improvement, from minimal to large improvement, across different activities.	We note that the Related Measure NQF #2613 also use these self-care data elements to measure improvement in self-care for the Skilled Nursing Facility setting.

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
4. Simplified and Targeted Rating Scale The function data elements used in this performance measure are coded using a 6-level rating scale that indicates the patient's level of independence performing an activity; higher scores indicate more independence. The decision to use a 6-level rating scale was based on several factors. First, input from the clinical communities and research examining the relationship between minutes of assistance and functional assessment scores. Second, scores do not decrease due to the use of an assistive device, which is consistent with the approach used by the World Health Organization's International Classification of Functioning (ICF) that suggests what matters most is someone's capacity to do an activity regardless of the use of assistive devices. Thus, the 6-level rating scale was designed to measure a person's ability to perform daily activities with or without assistive devices. The rating scale focused solely on the type and amount of human assistance needed to compete an activity.	Another measure of self-care function, Related Measure NQF #2613, used in the Skilled Nursing Facility, also adopted the 6-level rating scale.
5. Meaningful Activity Not Attempted Codes The use of four distinct activity not attempted codes were implemented so that providers code a specific reason for an activity not being attempted. For example, code 07 is used if the patient refused to attempt the self-care activity, such as putting on/taking off footwear, during the entire 3-day assessment period. If the patient was not able to perform the activity safely, due to medical or safety concerns, code 88 is used. A qualified clinician's assessment that a patient's medical condition contributes to their inability to safely put on and take off footwear means something different than a patient who is refusing to perform the activity, and the coding responses that allow for this distinction.	Other measures of self-care function, such as Related Measure NQF #2613 used in the Skilled Nursing Facility also adopted the activity not attempted codes.
Measure Calculation	

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
1. Difference Approach for Interpretability This measure calculates the risk-adjusted performance score using observed and expected scores. When observed and expected scores are compared, the difference between the two scores is calculated, and this difference approach represents an additive relationship (i.e., the observed change in function minus the expected changed in function, plus the national average). The choice between using a difference or a ratio approach depends on the researcher's assumption on whether the relationship between risk factors and the outcome is additive or multiplicative (Mukamel et al., 2000). After we conducted testing using the two approaches, and consulted with methodological experts, we decided to use the difference approach for this measure. When the expected value is small, the ratio is more volatile with small changes in the observed values (Ash et al, 2012). As the denominator approaches zero, the ratio can increase greatly in magnitude, as the observed values become greater than the expected values. Also, if the average expected value is 0, then the ratio cannot be calculated.	The following measures also use this approach: Related Measure NQF #2613, used in the Skilled Nursing Facility, and the FOTO measures (NQF# 0426, 0427, and 0428).

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
2. Exclusion Criteria to Maintain Validity We believe exclusion criteria are important specifications that support the validity of the quality measure. The exclusion criteria were selected with input from the Technical Expert Panel and input from a public comment process, as well as a review of existing literature. Patients with limited or less predictable self-care due to the nature of their medical condition improvement (e.g., severe brain damage) were recommended for exclusion by experts. Their reasoning was that attributing limited improvement in patients with these conditions to poor quality of care by the IRF would threaten the validity of the quality measure. The measure also has exclusions for patients with incomplete stays (e.g., discharged to acute care) or patients who were discharge to hospice for whom functional improvement may not be a goal.	The Related Measures NQF #2613 and #0688 also exclude patients with selected medical conditions where improvement is very unlikely in order to maintain the validity of the measures' performance scores. The Related Measures NQF #2613 and #0688 also exclude hospice patients from their performance measures.

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
3. Robust Risk Adjustment Model Improvement in functional abilities for patients in IRFs are associated with many patient demographic and clinical characteristics. Existing literature, stakeholder comments and technical expert opinions about risk adjustors were gathered and we all suggestions were tested with data. This measure adjusts for patient demographic and clinical characteristics, including age category, primary rehabilitation diagnosis, prior functioning, admission self-care or mobility functional status, cognitive function, communication function, and comorbidities. Adequate risk adjustment is critical to ensure quality measure validity, such that differences in performance scores across IRFs are related to differences in quality of care as much as possible, rather than to differences in patient characteristics across facilities. For an individual patient, up to 61 risk adjustors may apply in the self-care model. Notably, 40 of these are for comorbidities. This number of comorbidities are included in the model to account for differences in functional improvement for people with different co-existing health conditions. We would like to highlight that no patient in the national data had all 40 comorbidities and, in fact, the maximum number of comorbidities a person had was 11. On average, patients had only 1 comorbidity (mean = 1.4), and this means that the average patient has a "0" value for all other comorbidities in the model and a final risk adjustment model adjusting for 22 factors.	Because risk adjustment is imperative when measuring functional outcomes, the other measures such as the Competing Measure, NQF #2286 and the Related Measures #2613, #2769, #2775, #2776, #2777 and the FOTO measures (NQF# 0426, 0427, and 0428) also risk adjust for comorbidities.

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
4. Scale Construct Validity To ensure strong content and construct validity, the CMS self-care measures only include items related to the construct of self-care, as traditionally defined in functional assessment instruments. CMS recognizes that other aspects of functioning, such as cognition and communication, are important, however, data for these aspects of functioning are typically not aggregated with self-care data to measure improvement in self-care functioning. Existing literature supports the idea that cognition is a separate construct from motor function (i.e., mobility and self-care) when data from a diverse patient population are analyzed, and concludes that items related to mobility, self-care, and cognition should not be merged into a combined score (Avlund et al., 1993; Coster et al., 2004; Glenny & Stolee, 2009; Thomas et al., 1998). When selecting the data elements for this self-care measure, our goal was to measure self-care as precisely as possible, and therefore we did not to include items related to cognition.	
Feasibility, Usability and Use Considerations	

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
1. Use of Data Data are used by IRFs for internal quality improvement efforts (see "3. Confidential Feedback Reports" section below and are displayed to the public (see "4. Public Availability of Measure Data" below). The functional assessment data used to calculate this measure will be used by CMS to determine Prospective Payment rates for Medicare Part A patients treated in IRFs beginning in October 1, 2019. This data collected for quality measurement are also used for payment. There no costs associated with fees, licensing or other requirements associated with the measure data elements or risk model. All providers have access to a free Java-based software application to collect and maintain their facility's IRF-PAI information. Facilities are able to enter and subsequently export their data from the application for submission to the appropriate national data repository.	
2. Interpretability of Performance Score The performance measure score is presented publicly on IRF Compare as a mean risk-adjusted change in self-care score that is a continuous number and the typical method that IRFs report data. This makes the score more interpretable and transparent to stakeholders and end users. Feedback from Technical Experts in the development of the measures indicated their support for a summed raw item score with the importance of transparency of calculating the quality measure and the ease of data interpretation.	
3. Confidential Reports for Providers Free reports were made available to IRFs through the Certification and Survey Provider Enhanced Reports (CASPER) system starting in 2017. These reports contain feedback on providers' measure performance for internal quality improvement efforts and on national measure scores for quality benchmarking. More details about these reports and what measure data they contain is available in Section 4a2.1.2. under Usability and Use.	

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
4. Public Availability of Measure Data All measures reported in the IRF QRP serve two purposes: to reflect IRF provider performance by publicly disseminating data about quality of care, which help consumers' and family members' decision making, and to support providers in improving the quality of care they provide to patients. Public reporting on IRF Compare for the functional outcome measures will begin in fall 2020 (on discharges from January 1, 2019 through December 31, 2019). 5. Support for Interpretation and Calculation of Performance Scores To assist providers to collect accurate data for this measure, CMS has offered multiple in-person and on-line training opportunities since May 2015. In addition, several help desks are available to answer provider questions regarding data collection, and feedback reports, and "Q & A" documents are posted on the CMS website. To assist providers with calculating their facility's performance score internally, the publicly available IRF QRP Measure Calculations and Reporting User's Manual presents measure specifications and calculations for each measure included in the IRF QRP, including this measure. To assist consumers, such as family members and patients, with viewing and interpreting the measures posted on the public IRF Compare website, an IRF Public Reporting help desk is available. Individuals can submit questions or comments to CMS at any time and in this way, CMS provides real-time support to patients, families and caregivers seeking additional information or clarification on measures.	

References:

Ash, A.S., Shwartz M., Pekoz, E.A., & Hanchate, A.D. Chapter 12: Comparing outcomes across providers. In: Iezonni L, ed. Risk adjustment for measuring health care outcomes, 4th ed. Chicago, IL: Health administration Press; 2012:335-378.

Avlund K, Kreiner S, Schultz-Larsen K. Construct validation and the Rasch model: functional

ability of health elderly people. Scand J Soc Med. 1993;21:233–246.

Coster, W. J., Haley, S. M., Ludlow, L. H., Andres, P. L., & Ni, P. S. (2004). Development of an applied cognition scale to measure rehabilitation outcomes. Archives of physical medicine and rehabilitation, 85(12), 2030-2035.

Glenny, C., & Stolee, P. (2009). Comparing the functional independence measure and the interRAI/MDS for use in the functional assessment of older adults: a review of the literature. BMC geriatrics, 9(1), 52.

Mukamel, D. B., Dick, A., & Spector, W. D. (2000). Specification issues in measurement of quality of medical care using risk adjusted outcomes. Journal of Economic and Social Measurement, 26(3, 4), 267-281.

Thomas VS, Rockwood K, McDowell I. Multidimensionality in instrumental and basic activities of daily living. J Clin Epidemiol. 1998;51:315–321.

II. INPATIENT REHABILITATION FACILITY (IRF) FUNCTIONAL OUTCOME MEASURE: CHANGE IN MOBILITY SCORE FOR MEDICAL REHABILITATION PATIENTS (NQF #2634)

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0167: Improvement in Ambulation/locomotion

0175: Improvement in bed transferring

0422: Functional status change for patients with Knee impairments

0423: Functional status change for patients with Hip impairments

0424: Functional status change for patients with Foot and Ankle impairments

0425: Functional status change for patients with lumbar impairments

0426: Functional status change for patients with Shoulder impairments

0427: Functional status change for patients with elbow, wrist and hand impairments

0428: Functional status change for patients with General orthopaedic impairments

0688: Percent of Residents Whose Need for Help with Activities of Daily Living Has Increased (long stay)

2287: Functional Change: Change in Motor Score

2321: Functional Change: Change in Mobility Score

2612: CARE: Improvement in Mobility

2632: Long-Term Care Hospital (LTCH) Functional Outcome Measure: Change in Mobility Among Patients Requiring Ventilator Support

2643: Average change in functional status following lumbar spine fusion surgery

2653: Average change in functional status following total knee replacement surgery

2774: Functional Change: Change in Mobility Score for Skilled Nursing Facilities

2775: Functional Change: Change in Motor Score for Skilled Nursing Facilities

2776: Functional Change: Change in Motor Score in Long Term Acute Care Facilities

2778: Functional Change: Change in Mobility Score for Long Term Acute Care Facilities

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

All the listed measures address the same topic, function, but the target populations for most of these measures is not the IRF patient population. For example, measures are used for patients/residents treated in outpatient settings, home care, skilled nursing facilities, long-stay nursing homes, and long-term care hospitals. One measure has been previously identified by NQF staff as a competing measure: Functional Change: Change in Mobility Score (NQF #2321).

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

The NQF and the Patient Experience and Function Standing Committee may choose to endorse both competing measures, because both provide value. If NQF and the committee believe that only one measure should be endorsed as “best-in-class,” we offer a list of the strengths of our measure below as well as a comparison of feasibility, usability and use for consideration. Specifically, we describe the similarities and important differences between this change in mobility measure and the listed related and competing measures (See 5.1.a). We note that several features of this measure (e.g., the data elements, many of the risk adjustors, and the risk-adjustment approach) are the same as or aligned with the specifications of several of the other endorsed measures. Therefore, we believe that the specifications for this measure incorporate the best features of all endorsed related and competing measures, and, as a whole, represents the “best in class” for measuring change in mobility for IRFs.

This Change in Mobility (NQF #2634) measure was developed by building on the most recent science related to measurement of patient functioning and quality measure development. The latest science and scholarly literature, clinical thinking, and expert input on functional

assessment and quality measurement was combined with a cross-setting design and purpose in mind. Specifications were discussed with stakeholders and experts, pilot tested, and analyzed throughout the development process, as described in the Testing form.

Table 2. Features of the Measure and Notes about Harmonization: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients (NQF #2634)

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
Functional Assessment Data Elements	
1. Cross-Setting Design The functional assessment data elements for this measure, included in Section GG: Functional Abilities and Goals, were designed and tested with a cross-setting purpose in mind to ensure that data may be collected by clinicians in various post-acute and acute care settings. This enhances the cross-setting validity and reliability of quality measures that use these data. Standardization of self-care and mobility data elements across post-acute care settings has been an important goal for policymakers and included in the IMPACT Act of 2014.	We note that another measure focused on improvement in mobility, Related Measure NQF #2612, also use the data elements from Section GG: Functional Abilities and Goals as part of their performance measure with the rationale that the data elements were developed for cross-setting use and that the data elements are standardized.
2. Clinician Observation To determine a patient's functional ability, providers are instructed to code the data elements in Section GG: Functional Abilities and Goals primarily based on clinical observation. Specifically, a qualified clinician will assess the patient's performance based on direct observation, as well as gather input from reports from other clinicians, care staff, or family as well as the patient's self-report. Typically, an interdisciplinary team of qualified clinicians is involved in assessing the patient and CMS provides guidance through manuals, training programs, and help desk responses to support providers in collecting accurate functional assessment data.	We note that the Competing Measure NQF #2321 and Related Measures NQF #2612, #2774, #2775, #2776, and #2778 also use clinician observation to assess and code a patient's functional abilities.

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
3. Functional Assessment Data Elements Capture Range of Functioning The functional assessment data elements and associated rating scale were designed to build on the existing science of functional assessment, which included a review of the strengths and limitations of existing instruments. The inclusion of 15 mobility data elements allows for the measurement of a wide range of patient functioning and thus the opportunity to demonstrate gains in a variety of functional activities. Patients may be expected to make varying amounts of improvement, from minimal to large improvement, across different activities.	We note that the Related Measure NQF #2612 also use these mobility data elements to measure improvement in mobility for the Skilled Nursing Facility setting.
4. Simplified and Targeted Rating Scale The function data elements used in this performance measure are coded using a 6-level rating scale that indicates the patient's level of independence performing an activity; higher scores indicate more independence. The decision to use a 6-level rating scale was based on several factors. First, input from the clinical communities and research examining the relationship between minutes of assistance and functional assessment scores, which is curvilinear, indicated that persons with high functional assessment scores frequently did not require daily assistance. Second, scores do not decrease due to the use of an assistive device, which is consistent with the approach used by the World Health Organization's International Classification of Functioning (ICF) that suggests what matters most is someone's capacity to do an activity regardless of the use of assistive devices. Thus, the 6-level rating scale was designed to measure a person's ability to perform daily activities with or without assistive devices. The rating scale focused solely on the type and amount of human assistance needed to compete an activity.	Another measure of mobility function, Related Measure NQF #2612 used in the Skilled Nursing Facility, also adopted the 6-level rating scale.

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
5. Meaningful Activity Not Attempted Codes The use of four distinct activity not attempted codes were implemented so that providers to code a specific reason for an activity not being attempted. For example, code 07 is used if the patient refused to attempt the mobility activity, such as walking 150 feet, during the entire 3-day assessment period. If the patient was not able to perform the activity safely, due to medical or safety concerns, code 88 is used. A qualified clinician's assessment that a patient's medical condition contributes to their inability to safely walk 150 feet means something different than a patient who is refusing to perform the activity, and the coding responses that allow for this distinction.	Other measures of mobility function, such as Related Measure NQF #2612 used in the Skilled Nursing Facility, also adopted the activity not attempted codes.
Measure Calculation	
1. Difference Approach for Interpretability This measure calculates the risk-adjusted performance score using observed and expected scores. When observed and expected scores are compared, the difference between the two scores is calculated, and this difference approach represents an additive relationship (i.e., the observed change in function minus the expected changed in function, plus the national average). The choice between using a difference or a ratio approach depends on the researcher's assumption on whether the relationship between risk factors and the outcome is additive or multiplicative (Mukamel et al., 2000). After we conducted testing using the two approaches, and consulted with methodological experts, we decided to use the difference approach for this measure. When the expected value is small, the ratio is more volatile with small changes in the observed values (Ash et al, 2003). As the denominator approaches zero, the ratio can increase greatly in magnitude, as the observed values become greater than the expected values. Also, if the average expected value is 0, then the ratio cannot be calculated.	The following measures also use this approach: Related Measure NQF #2612, used in the Skilled Nursing Facility, and the FOTO measures (NQF #0422, 0423, 0424, 0425, 0426, 0427, and 0428).

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
2. Exclusion Criteria to Maintain Validity We believe exclusion criteria are important specification that support the validity of the quality measure. The exclusion criteria were selected with input from the Technical Expert Panel and input from a public comment process, as well as a review of existing literature. Patients with limited or less predictable mobility due to the nature of their medical condition improvement (e.g., severe brain damage) were recommended for exclusion by experts. Their reasoning was that attributing limited improvement in patients with these conditions to poor quality of care by the IRF would threaten the validity of the quality measure. The change in mobility measure also has exclusions for patients with incomplete stays (e.g., discharged to acute care) or patients who were discharge to hospice for whom functional improvement may not be a goal.	The Related Measures NQF #2612 and #2643 also exclude patients with selected medical conditions where improvement is very unlikely or unexpected in order to maintain the validity of the measure's performance score. The Related Measures NQF #2612 and #0688 also exclude hospice patients from their performance measure.

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
3. Robust Risk Adjustment Model Improvement in functional abilities for patients in IRFs are associated with many patient demographic and clinical characteristics. Existing literature, stakeholder comments and technical expert opinions about risk adjustors were gathered and we all suggestions were tested with data. This measure adjusts for patient demographic and clinical characteristics, including age category, primary rehabilitation diagnosis, prior functioning, admission self-care or mobility functional status, cognitive function, communication function, and comorbidities. Adequate risk adjustment is critical to ensure quality measure validity, such that differences in performance scores across IRFs are related to differences in quality of care as much as possible, rather than to differences in patient characteristics across facilities. For an individual patient, up to 72 risk adjustors may apply in the mobility model. Notably, 50 of these are for comorbidities. This number of comorbidities are included in the model to account for differences in functional improvement for people with different co-existing health conditions. We would like to highlight that no patient in the national data had all 50 comorbidities and, in fact, the maximum number of comorbidities a person had was 10. On average, patients had only 2 comorbidities (mean = 1.6), and this means that the average patient has a "0" value for all other comorbidities in the model and a final risk adjustment model adjusting for 24 factors.	Because risk adjustment is imperative when measuring functional outcomes, the other measures such as the Competing Measure, NQF #2321 and Related Measures such as #2612, #2774, #2778 and the FOTO measures (NQF #0422, 0423, 0424, 0425, 0426, 0427, and 0428) also risk adjust for comorbidities.
4. Inclusion of Patients Who Use a Wheelchair The CMS mobility measures include the wheelchair mobility items as part of the performance measure to reflect that some patients use a wheelchair as their primary method of mobility.	We note that the Competing Measure NQF #2321 as well as the Related Measures #2612, #2774 and #2778 also include wheelchair mobility in their quality measure calculation.
Feasibility, Usability and Use Considerations	

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
1. Use of Data Data are used by IRFs for internal quality improvement efforts (see "3. Confidential Feedback Reports" section below and are displayed to the public (see "4. Public Availability of Measure Data" below). The functional assessment data used to calculate this measure will be used by CMS to determine Prospective Payment rates for Medicare part A patients treated in IRFs beginning in October 1, 2019. This data collected for quality measurement are also used for payment. There no costs associated with fees, licensing or other requirements associated with the measure data elements or risk model. All providers have access to a free Java-based software application to collect and maintain their facility's IRF-PAI information. Facilities are able to enter and subsequently export their data from the application for submission to the appropriate national data repository.	
2. Interpretability of Performance Score The performance measure score is presented publicly on IRF Compare as a mean change in mobility score that is a continuous number and the typical method that IRFs report data. This makes the score more interpretable and transparent to stakeholders and end users. Feedback from Technical Experts in the development of the measures indicated their support for a summed raw item score with the importance of transparency of calculating the quality measure and the ease of data interpretation.	
3. Confidential Reports for Providers Free reports were made available to IRFs through the Certification and Survey Provider Enhanced Reports (CASPER) system starting in 2017. These reports contain feedback on providers' measure performance for internal quality improvement efforts and on national measure scores for quality benchmarking. More details about these reports and what measure data they contain is available in Section 4a2.1.2. under Usability and Use.	

FEATURES OF MEASURE	NOTES ABOUT HARMONIZATION
4. Public Availability of Measure Data All measures reported in the IRF QRP serve two purposes: to reflect IRF provider performance by publicly disseminating data about quality of care, which help consumers' and family members' decision making, and to support providers in improving the quality of care they provide to patients. Public reporting on IRF Compare for the functional outcome measures will begin in fall 2020 (on discharges from January 1, 2019 through December 31, 2019).	

References:

Ash, A.S., Shwartz M., Pekoz, E.A., & Hanchate, A.D. Chapter 12: Comparing outcomes across providers. In: Iezonni L, ed. Risk adjustment for measuring health care outcomes, 4th ed. Chicago, IL: Health administration Press; 2012:335-378.

Avlund K, Kreiner S, Schultz-Larsen K. Construct validation and the Rasch model: functional ability of health elderly people. Scand J Soc Med. 1993;21:233–246.

Coster, W. J., Haley, S. M., Ludlow, L. H., Andres, P. L., & Ni, P. S. (2004). Development of an applied cognition scale to measure rehabilitation outcomes. Archives of physical medicine and rehabilitation, 85(12), 2030-2035.

Glenny, C., & Stolee, P. (2009). Comparing the functional independence measure and the interRAI/MDS for use in the functional assessment of older adults: a review of the literature. BMC geriatrics, 9(1), 52.

Mukamel, D. B., Dick, A., & Spector, W. D. (2000). Specification issues in measurement of quality of medical care using risk adjusted outcomes. Journal of Economic and Social Measurement, 26(3, 4), 267-281.

Thomas VS, Rockwood K, McDowell I. Multidimensionality in instrumental and basic activities of daily living. J Clin Epidemiol. 1998;51:315–321.

Appendix B1: Competing Measures (tabular format)

	2286 Functional Change: Change in Self Care Score	2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients
Steward	Uniform Data System for Medical Rehabilitation	Centers for Medicare & Medicaid Services
Description	Change in rasch derived values of self-care function from admission to discharge among adult patients treated at an inpatient rehabilitation facility who were discharged alive. The timeframe for the measure is 12 months. The measure includes the following 8 items: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory.	This measure estimates the risk-adjusted mean change in self-care score between admission and discharge for Inpatient Rehabilitation Facility (IRF) Medicare patients.
Type	Outcome	Outcome
Data Source	Instrument-Based Data, Other	Instrument-Based Data
Level	Facility, Other	Facility
Setting	Inpatient/Hospital	Post-Acute Care
Numerator Statement	Average change in rasch derived self-care functional score from admission to discharge at the facility level, including items: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory. Average is calculated as: (sum of change at the patient level for all items (Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory) / total number of patients).	The measure does not have a simple form for the numerator and denominator. This measure estimates the risk-adjusted change in self-care score between admission and discharge among Inpatient Rehabilitation Facility (IRF) Medicare patients age 21 or older. The change in self-care score is calculated as the difference between the discharge self-care score and the admission self-care score.
Numerator Details	For Inpatient Rehabilitation Facilities (IRFs) data collection currently occurs as required by the Centers for Medicare and Medicaid Services (CMS) reimbursement using the mandated payment document, the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI). Embedded in the IRF-PAI is the FIM® Instrument. The FIM® Instrument is a criterion referenced tool with 18 items that measure patient physical and cognitive functional status and patient burden of care (level of dependence/need for helper assistance). Each item is rated on a scale of 1 (most dependent) to 7 (completely independent). For the purposes of this measure, a subset of 8 FIM® items has been tested and validated which comprise the self-care measure; those items are: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory. Rasch analysis was performed on the 8 items and the difference in the rasch derived values (defined in S.2b) from admission to discharge reflect the change at the patient level. The numerator of the measure is the facility´s average change. While the IRF-PAI is specific to inpatient rehabilitation facilities, the measure can be used in all post-acute care venues. The FIM® instrument can be assessed in all venues of care and has been tested and validated for use in inpatient medical rehabilitation, long term acute care facilities (LTAC), skilled nursing facilities (SNF) and home health. At present, numerous LTACs and SNFs utilize the FIM® instrument (www.udsmr.org), thus the self-care measure is applicable for use in IRF, SNF, LTAC and other venues where patient functional change is anticipated.	Seven self-care activities are each scored based on a patient´s ability to complete the activity. The scores for the seven activities are summed to obtain a self-care score at the time of admission and at the time of discharge. The change in self-care is the difference between the discharge self-care score and the admission self-care score. The 7 self-care items are: GG0130A. Eating GG0130B. Oral hygiene GG0130C. Toileting hygiene GG0130E. Shower/bathe self GG0130F. Upper body dressing GG0130G. Lower body dressing GG0130H. Putting on/taking off footwear Each patient´s ability to complete each self-care activity (item) is rated by a clinician using the following 6-level rating scale: level 06 - Independent level 05 - Setup or clean up assistance level 04 - Supervision or touching assistance level 03 - Partial/moderate assistance level 02 - Substantial/maximal assistance level 01 - Dependent If the patient did not attempt the activity, the reason that the activity did not occur is reported as: 07 = Patient refused 09 = Not applicable 10 = Not attempted due to environmental limitations 88 = Not attempted due to medical condition or safety concerns. The performance period is 12 months for reporting on CMS's IRF Compare website.
Denominator Statement	Facility adjusted adjusted expected change in rasch derived values, adjusted at the Case Mix Group level.	The denominator is the number of Inpatient Rehabilitation Facility Medicare patient stays, except those that meet the exclusion criteria.
Denominator Details	To calculate the facility adjusted expected self-care change in rasch derived values, indirect standardization is used, which weights national CMG-specific values by facility-specific CMG proportions. CMG-adjustment derives the expected value based on the case mix and severity mix of each facility. The case-mix group (CMG) classification system groups similarly impaired patients based on functional status at admission (in essence, patient severity). Patients within the same CMG are expected to have similar resource utilization needs and similar outcomes. There are three steps to classifying a patient into a CMG at admission: 1. Identify the patient's impairment group code (IGC). 2. Calculate the patient's weighted motor index score, calculated from 12 of the 13 motor FIM® items as indicated on the CMS IRF-PAI v. 20 instrument (attached). 3. Calculate the cognitive FIM® rating (as indicated on the CMS IRF_PAI v. 20 instrument) and the patient age at admission. (This step is not required for all CMGs.) See file uploaded in S.15 for calculations. While CMGs are only present for patients admitted to an IRF, the same procedure can be used for patients receiving care at a LTAC facility and/or a SNF, with groupings specific to those venues of care.	The denominator is the number of Inpatient Rehabilitation Facility Medicare Part A and Medicare Advantage patient stays, except those that meet the exclusion criteria.
Exclusions	National values used in the CMG-adjustment procedure will not include cases who died in the IRF or cases less than 18 years old. It is standard to exclude cases who died during rehabilitation as this is a highly atypical outcome, in addition, minors are excluded as well. The measure testing file includes further explanation regarding the exclusion criteria as well as references.	This quality measure has six patient-level exclusion criteria: 1) Patients with incomplete stays. Rationale: When a patient has an incomplete stay, for example, the patients leave urgently due to a medical emergency, it can be challenging to gather accurate discharge functional status data. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical

	2286 Functional Change: Change in Self Care Score	2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients
		advice; and patients with a length of stay less than 3 days. 2) Patients who are independent with all self-care activities at the time of admission. Rationale: Patients who are independent with all the self-care items at the time of admission are assigned the highest score on all the self-care items, and thus, would not be able to show functional improvement on this same set of items at discharge. 3) Patients with the following medical conditions on admission: coma; persistent vegetative state; complete quadriplegia; locked-in syndrome; or severe anoxic brain damage, cerebral edema or compression of the brain. Rationale: These patients are excluded because they may have limited or less predictable self-care improvement with the selected self-care items. 4) Patients younger than age 21. Rationale: There is only limited evidence published about functional outcomes for individuals with Medicare who are younger than 21. 5) Patients discharged to Hospice. Rationale: Patient goals may change during the IRF stay, and functional improvement may no longer be a goal for a patient discharged to hospice. 6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries. Rationale: IRF-PAI data for patients not covered by the Medicare program are not submitted to the Centers for Medicare and Medicaid Services. Facility-level quality measure exclusion: For IRFs with fewer than 20 patient stays, data for this quality measure are not publicly reported.
Exclusion Details	Patient date of birth (DOB) and discharge setting are both variables collected in the IRF-PAI. Age can be calculated from DOB and patient date of admission (also collected in the IRF-PAI). In the variable discharge setting, there is a specific category for ´died´ (code: 11). Date of birth, date of admission and discharge setting (including died as a category) are also assessed in the LTAC and SNF.	The following items are used to identify which patients are excluded from the quality measure calculations. These data elements are included on the current version of the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), which is available at: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/IRFPAI.html It can be challenging to gather accurate discharge functional status data for patients who experience incomplete stays. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients discharged to a hospice; patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days. Items used to identify these patient records: 1) Patients with incomplete stays. Patients with a length of stay less than 3 days: We calculate length of stay using the following items on the IRF-PAI. Item 12. Admission Date. Item 40. Discharge Date. Length of stay is calculated as the Discharge Date minus the Admission Date (Discharge Date - Admission Date). Patient records with a length of stay of less than 3 days are excluded. Item 41. Patient discharged against medical advice. This item is used to identify patients discharged against medical advice. Patient records with a response of “Yes = 1" are excluded. Item 44C. Was the patient discharged alive? This item is used to identify patients who died during the IRF stay. Patient records with a response of “No=0" are excluded. 44D. Patient’s discharge destination/living setting. This item is used to identify patients with an incomplete stay. Short-term General Hospital = 02 Long-Term Care Hospital = 63 Inpatient Psychiatric Facility = 65 Critical Access Hospital = 66. 2) Patients who are independent with all self-care activities at the time of admission: Patients who are independent with all the self-care items at the time of admission are assigned the highest score on all the self-care items, and thus, would not be able to show functional improvement (i.e., a higher score)on this same set of items at discharge. Self-care items GG0130A. Eating = 06, and GG0130B. Oral hygiene = 06, and GG0130C. Toileting hygiene = 06, and GG0130E. Shower/bathe self = 06, and GG0130F. Upper body dressing = 06, and GG0130G. Lower body dressing = 06, and GG0130H. Putting on/taking off footwear = 06. 3) Patients with the following medical conditions on admission: coma; persistent vegetative state; complete quadriplegia; and locked-in syndrome; and severe anoxic brain damage, cerebral edema or compression of the brain. The following items will be used to identify patients with these conditions: 21A. Impairment Group. 0004.1221 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia

	2286 Functional Change: Change in Self Care Score	2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients
		Complete, C1-C4 0004.1222 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia Complete, C5-C8 0004.2221 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C1-C4 0004.2222 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C5-C8 22. Etiologic Diagnosis. This item is used to determine a patient´s etiologic problem that led to the condition for which the patient is receiving rehabilitation. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to identify and exclude the records of patients with these conditions: HCC 80. Coma, Brain Compression/Anoxic Damage ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela ICD-10-CM. G83.5. Locked-in state 24. Comorbid Conditions. This item is used to exclude selected comorbidities. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to identify and exclude the records of patients with these conditions: HCC 80. Coma, Brain Compression/Anoxic Damage ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela ICD-10-CM. G83.5. Locked-in state 4) Patients younger than age 21. These items are used to calculate age, and patients who are younger than 21 years of age at the time of admission are excluded. 6. Birth Date 12. Admission Date Age is calculated as the Admission Date minus the Birth Date (Admission Date - Birth Date). Patients younger than 21 are excluded. 5) Patients discharged to hospice 44D. Patient's discharge destination/living setting. This item is used to identify patients discharged to hospice. The following responses are used: Hospice (home) = 50 Hospice (institutional facility) = 51 6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries. The following items are used to identify and exclude the records of patients who are not Medicare Part A and Medicare Advantage beneficiaries: 20A. Primary Source = 99 - Not Listed AND 20B. Secondary Source = 99 - Not Listed
Risk Adjustment	Stratification by risk category/subgroup	Statistical Risk Model
Stratification	While the measure can be stratified by specific impairment type (using IGC), the CMG adjustment procedure allows for the measure to be complete, accurate, and valid for all patients within the facility, excluding cases who died and excluding patient under age 18 years.	N/A
Type Score	Ratio	Continuous variable, e.g. average
Algorithm	1. Target population: Inpatient rehabilitation facility patients, skilled nursing facility short term patients, long term acute care facility patients, and home health patients. 2. Exclusions: Age less than 18 and cases who died during the episode of care. 3. Cases meeting target process: All remaining cases. 4. Outcome: Ratio of facility level average self-care change (rasch derived values) to facility CMG adjusted expected self-care change. 5. Risk adjustment: CMG adjustment using indirect standardization of the proportion of cases at the facility by CMG, and CMG specific national average of rasch derived value of self-care change.	We provide the detailed calculation algorithm in an attachment entitled "IRF Detailed Function QM Specifications 2633 01-07-2019" included in the Appendix. The detailed calculation algorithm is provided to the public in the document entitled IRF Measure Calculations and Reporting User's Manual. The current version of this document is available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/IRF-Quality-Reporting-Program-Measures-Information-.html The following are the key steps used to calculate the measure: 1) Sum the scores of the admission self-care items to create an admission self-care score for each patient, after 'activity not attempted' codes (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are recoded. (range: 7 to 42). 2) Sum the scores of the discharge self-care items to create a discharge self-care score for each patient, after 'activity not attempted' codes (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are

	2286 Functional Change: Change in Self Care Score	2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients
		recoded. (range: 7 to 42). 3) Identify the records of patients who meet the exclusion criteria and exclude them from analyses. 4) Calculate the difference between the admission self-care score (from step 1) and the discharge self-care score (from step 2) for each patient to create a change in self-care score for each patient. 5) Calculate an expected change in self-care score for each patient using regression coefficients from national data and each patient's admission characteristics (risk adjustors). 6) Calculate an average change in self-care score for each IRF. This is the facility-level observed change in self-care score. 7) Calculate an average expected change in self-care score for each IRF. This is the facility-level expected change in self-care score. 8) Subtract the facility-level expected change score from the facility -level observed change score to determine the difference in scores (difference value). A difference value that is 0 indicates the observed and expected scores are equal. An observed minus expected difference value that is higher than 0 (positive value) indicates that the observed change score is greater (better) than the expected change score. An observed minus expected difference value that is less than 0 (negative value) indicates that the observed change score is lower (worse) than the expected change score. 9) Add the national average change in self-care score to each IRF's difference value (from step 8). This is the risk-adjusted mean change in self-care score. Each patient´s ability to complete each self-care activity (item) is rated by a clinician using the following 6-level rating scale: level 06 - Independent level 05 - Setup or clean up assistance level 04 - Supervision or touching assistance level 03 - Partial/moderate assistance level 02 - Substantial/maximal assistance level 01 - Dependent The 7 self-care items are: GG0130A. Eating GG0130B. Oral hygiene GG0130C. Toileting hygiene GG0130E. Shower/bathe self GG0130F. Upper body dressing GG0130G. Lower body dressing GG0130H. Putting on/taking off footwear

	2321 Functional Change: Change in Mobility Score	2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients
Steward	Uniform Data System for Medical Rehabilitation	Centers for Medicare & Medicaid Services
Description	Change in rasch derived values of mobility function from admission to discharge among adult inpatient rehabilitation facility patients aged 18 years and older who were discharged alive. The timeframe for the measure is 12 months. The measure includes the following 4 mobility FIM® items:Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs.	This measure estimates the mean risk-adjusted mean change in mobility score between admission and discharge for Inpatient Rehabilitation Facility (IRF) Medicare Part A and Medicare Advantage patients.
Type	Outcome	Outcome
Data Source	Instrument-Based Data, Other	Instrument-Based Data
Level	Facility, Other	Facility
Setting	Inpatient/Hospital, Post-Acute Care	Post-Acute Care
Numerator Statement	Average change in rasch derived mobility function score from admission to discharge at the facility level. Includes the following items: Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs. Average is calculated as: (sum of change at the patient level/total number of patients). Patient less than 18 years of age at admission to the facility or patients who died within the facility are excluded.	The measure does not have a simple form for the numerator and denominator. This measure estimates the risk-adjusted change in mobility score between admission and discharge among Inpatient Rehabilitation Facility (IRF) Medicare Part A and Medicare Advantage patients age 21 and older. The change in mobility score is calculated as the difference between the discharge mobility score and the admission mobility score.
Numerator Details	For Inpatient Rehabilitation Facilities (IRFs) data collection is presently required for payment reimbursement by the Centers for Medicare and Medicaid Services (CMS) using the mandated Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI). Embedded in the IRF-PAI is the FIM® Instrument. The FIM® Instrument is a criterion referenced tool with 18 items that measures patient physical and cognitive function, need for helper assistance, burden of care/level of dependence. Each item is rated on a scale of 1 (most dependent) to 7 (completely independent). For the purposes of this measure, a subset of 4 FIM® items has been tested and validated as the Change in Mobility measure; the items are: Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs. Rasch analysis was performed on the items and the difference in the rasch derived values (defined in S.2b) from admission to discharge reflect the change at the patient level. The numerator of the measure is the average change in mobility score at the facility level. While the IRF-PAI is specific to inpatient rehabilitation facilities, the change in mobility measure can be used in all post-acute care venues. The FIM® instrument is routinely used for patient functional assessment in all venues of care and has been tested and validated for use in IRFs, skilled nursing facilities (SNFs) and long term acute care facilities (LTAC) (www.udsmr.org), therefore this measure is not specific for inpatient medical rehabilitation use only.	Seventeen mobility activities are each scored based on a patient´s ability to complete the activity. The scores for the activities are summed to obtain a mobility score at the time of admission and at the time of discharge. The change in mobility is the difference between the discharge mobility score and the admission mobility score. The mobility items are: GG0170A. Roll left and right GG0170B. Sit to lying GG0170C. Lying to sitting on side of bed GG0170D. Sit to stand GG0170E. Chair/bed-to-chair transfer GG0170F. Toilet transfer GG0170G. Car transfer GG0170I. Walk 10 feet GG0170J. Walk 50 feet with two turns GG0170K. Walk 150 feet GG0170L. Walking 10 feet on uneven surfaces GG1070M. 1 step (curb) GG0170N. 4 steps GG0170O. 12 steps GG0170P. Picking up object GG0170R. Wheel 50 feet with two turns (for patients who do not walk at admission and discharge) GG0170S. Wheel 150 feet (for patients who do not walk at admission and discharge) Each patient´s ability to complete each mobility activity (item) is rated by a clinician using the following 6-level rating scale: level 06 - Independent level 05 - Setup or clean up assistance level 04 - Supervision or touching assistance level 03 - Partial/moderate assistance level 02 - Substantial/maximal assistance level 01 - Dependent If the patient did not attempt the activity, the reason that activity did not occur is reported as: 07 = Patient refused 09 = Not applicable 10 = Not attempted due to environmental limitations 88 = Not attempted due to medical condition or safety concerns. The performance period is 12 months for reporting on CMS's IRF Compare website.
Denominator Statement	Facility adjusted expected change in rasch derived mobility values, adjusted at the Case Mix Group (CMG) level.	The denominator is the number of Inpatient Rehabilitation Facility Medicare Part A and Medicare Advantage patient stays, except those that meet the exclusion criteria.
Denominator Details	To calculate the facility adjusted expected change in rasch derived mobility values, indirect standardization was used, which weights national CMG-specific values by facility-specific CMG proportions. CMG-adjustment derives the expected value based on the case mix and severity mix of each facility. The case-mix group (CMG) classification system groups similarly impaired patients based on functional status at admission, in essence, patient severity. Patients within the same CMG are expected to have similar resource utilization needs and similar functional outcomes. There are three steps to classifying a patient into a CMG at admission: 1. Identify the patient's impairment group code (IGC). 2. Calculate the patient's weighted motor index score, calculated from 12 of the 13 motor FIM® items. 3. Calculate the cognitive FIM® rating and the patient´s age at admission. (This step is not required for all CMGs.)	The denominator is the number of Inpatient Rehabilitation Facility Medicare Part A and Medicare Advantage patient stays, except those that meet the exclusion criteria.

	2321 Functional Change: Change in Mobility Score	2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients
	See file uploaded in S.2b for calculations or ´CMG Version 3.00 [ZIP, 9.02mb]´ at the following link for more details: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/CMG.html	
Exclusions	National values used in the CMG adjustment procedure will not include cases who died in the IRF or patients less than 18 years of age at admission. Cases who died during rehabilitation are not typical patients and are routinely omitted from reports and published research on rehabilitation outcomes. Further details and references related to the exclusion criteria can be found in the Measure Testing form.	This quality measure has six patient-level exclusion criteria: 1) Patients with incomplete stays. Rationale: When a patient has an incomplete stay, for example, the patients leave urgently due to a medical emergency, it can be challenging to gather accurate discharge functional status data. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days. 2) Patients who are independent with all mobility activities at the time of admission. Rationale: Patients who are independent with all the mobility items (with the exception of the wheelchair items GG0170R and GG0170S) at the time of admission are assigned the highest score on all the mobility items, and thus, would not be able to show functional improvement on this same set of items at discharge. 3) Patients with the following medical conditions on admission: coma, persistent vegetative state; complete quadriplegia; locked-in syndrome or severe anoxic brain damage, cerebral edema or compression of brain. Rationale: These patients are excluded because they may have limited or less predictable mobility improvement with the selected mobility items. 4) Patients younger than age 21. Rationale: There is only limited evidence published about functional outcomes for individuals with Medicare who are younger than 21. 5) Patients discharged to hospice. Rationale: Patient goals may change during the IRF stay, and functional improvement may no longer be a goal for a patient discharged to hospice. 6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries. Rationale: IRF-PAI data for patients not covered by the Medicare program are not submitted to the Centers for Medicare and Medicaid Services. Facility-level quality measure exclusion: For IRFs with fewer than 20 patient stays, data for this quality measure are not publicly reported.
Exclusion Details	Patient date of birth (DOB), date of admission and discharge setting variables are collected in the IRF-PAI. Age can be calculated from DOB and admission date. The variable discharge setting includes a category for ´died´ which is indicated as a code of ´11´. Patient date of birth, admission date and discharge setting are also documented in SNFs and LTAC facilities.	The following items are used to identify which patients are excluded from the quality measure calculations. These data elements are included on the current version of the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), which is available at: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/IRFPAI.html It can be challenging to gather accurate discharge functional status data for patients who experience incomplete stays. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients discharged to a hospice; patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days. Items used to identify these patient records: 1) Patients with incomplete stays. Patients with a length of stay less than 3 days: We calculate length of stay using the following items on the IRF-PAI. Length of stay is calculated as the Discharge Date minus the Admission Date (Discharge Date - Admission Date). Patient records with a length of stay of less than 3 days are excluded. Item 12. Admission Date. Item 40. Discharge Date. Item 41. Patient discharged against medical advice. This item is used to identify patients discharged against medical advice. Patient records with a response of "Yes = 1" are excluded. Item 44C. Was the patient discharged alive? This item is used to identify patients who died during the IRF stay. Patient records with a response of "No = 0" are excluded. 44D. Patient’s discharge destination/living setting. This item is used to identify an incomplete stay. Specifically, the following responses will be used to identity patients with incomplete stays: Short-term General Hospital = 02 Long-Term Care Hospital = 63 Inpatient Psychiatric Facility = 65 Critical Access Hospital = 66.

	2321 Functional Change: Change in Mobility Score	2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients
		2) Patients who are independent with all mobility activities at the time of admission. Patients who are independent with all the mobility items at the time of admission are assigned the highest score on all the mobility items, thus, would not be able to show functional improvement (i.e., a higher score) on this same set of items at discharge. The following items and scores are used to identify and exclude patient records: Mobility items GG0170A. Roll left and right = 06, and GG0170B. Sit to lying = 06, and GG0170C. Lying to sitting on side of bed = 06, and GG0170D. Sit to stand = 06, and GG0170E. Chair/bed-to-chair transfer = 06, and GG0170F. Toilet transfer = 06, and GG0170G. Car transfer = 06, and GG0170I. Walk 10 feet = 06, and GG0170J. Walk 50 feet with two turns = 06, and GG0170K. Walk 150 feet = 06, and GG0170L. Walking 10 feet on uneven surfaces = 06, and GG0170M. 1 step (curb) = 06, and GG0170N. 4 steps = 06, and GG0170O. 12 steps = 06, and GG0170P. Picking up object = 06. 3) Patients with the following medical conditions on admission: coma; persistent vegetative state; complete quadriplegia; locked-in syndrome; and severe anoxic brain damage, cerebral edema or compression of the brain. The following items will be used to identify patients with these conditions: 21A. Impairment Group. 0004.1221 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia Complete, C1-C4 0004.1222 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia Complete, C5-C8 0004.2221 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C1-C4 0004.2222 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C5-C8 22. Etiologic Diagnosis. This item is used to determine a patient´s etiologic problem that led to the condition for which the patient is receiving rehabilitation. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to identify and exclude records of patients with these conditions: HCC 80. Coma, Brain Compression/Anoxic Damage ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela ICD-10-CM. G83.5. Locked-in state 24. Comorbid Conditions. This item is used to exclude selected comorbidities. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to exclude records of patients with these conditions: HCC 80. Coma, Brain Compression/Anoxic Damage ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela ICD-10-CM. G83.5. Locked-in state 4) Patients younger than age 21. These items are used to calculate age, and patients who are younger than 21 years of age at the time of admission are excluded. 6. Birth Date 12. Admission Date Age is calculated as the Admission Date minus the Birth Date (Admission Date - Birth Date). Patients younger than 21 are excluded. 5) Patients discharged to hospice. 44D. Patient's discharge destination/living setting. This item is used to identify patients discharged to hospice. The following responses are used:

	2321 Functional Change: Change in Mobility Score	2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients
		Hospice (home) = 50 Hospice (institutional facility) = 51 6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries 20A. Primary Source = 99 - Not Listed AND 20B. Secondary Source = 99 - Not Listed
Risk Adjustment	Stratification by risk category/subgroup	Statistical Risk Model
Stratification	While the measure can be stratified by specific impairment type (IGC), the CMG adjustment procedure allows for the measure to be complete, accurate, and valid for all patients within the facility, excluding died cases and ages less than 18.	N/A
Type Score	Ratio	Continuous variable, e.g. average
Algorithm	1. Target population: patients receiving care at an inpatient medical rehabilitation facility, a skilled nursing facility, or a long term acute care facility. 2. Exclusions: Age less than 18 years and patients who died during the episode of care. 3. Cases meeting target process: All remaining cases. 4. Outcome: Ratio of facility level average mobility change (rasch derived values) to facility CMG adjusted expected mobility change. 5. Risk adjustment: CMG adjustment using indirect standardization of the proportion of cases at the facility by CMG, and CMG specific national average of rasch derived value of mobility change.	We provide the detailed calculation algorithm in an attachment entitled "IRF Detailed Function QM Specifications 2634 01-07-2019" included in the Appendix. The detailed calculation algorithm is provided to the public in the document entitled IRF Measure Calculations and Reporting User's Manual. The current version of this document is available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/IRF-Quality-Reporting-Program-Measures-Information-.html The following are key steps used to calculate the measure: 1) Sum the scores of the admission mobility items to create an admission mobility score for each patient, after 'activity not attempted' codes (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are recoded, and for patients who do not walk on admission and discharge, walking items have been recoded to use wheelchair mobility item codes. (range: 15 to 90). 2) Sum the scores of the discharge mobility items to create a discharge mobility score for each patient, after 'activity not attempted' values (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are recoded. As described in step 1, for patients who do not walk on admission and discharge, use wheelchair mobility item codes instead of walking codes. (range: 15 to 90). 3) Identify the records of patients who meet the exclusion criteria and exclude them from analyses. 4) Calculate the difference between the admission mobility score (from step 1) and the discharge mobility score (from step 2) for each patient to create a change in mobility score for each patient. 5) Calculate an expected change in mobility score for each patient using regression coefficients from national data and each patient's admission characteristics (risk adjustors). 6) Calculate an average observed change in mobility score for each IRF (using the patient data calculated in step 4). This is the facility-level observed change in mobility score. 7) Calculate an average expected change in mobility score for each IRF (using the patient data from step 5). This is the facility-level expected change in mobility score. 8) Subtract the facility-level expected change score from the facility-level observed change score to determine the difference in scores (difference value). A difference value that is 0 indicates the observed and expected scores are equal. An observed minus expected difference value that is higher than 0 (positive) indicates that the observed change score is greater (better) than the expected change score. An observed minus expected difference value that is less than 0 (negative) indicates that the observed change score is lower (worse) than the expected change score. 9) Add the national average change in mobility score to each IRF's difference value (from step 8). This is the risk-adjusted mean change in mobility score. Each patient´s ability to complete each mobility activity (item) is rated by a clinician using the following 6-level rating scale: level 06 - Independent level 05 - Setup or clean up assistance level 04 - Supervision or touching assistance level 03 - Partial/moderate assistance level 02 - Substantial/maximal assistance level 01 - Dependent The mobility items are: GG0170A. Roll left and right GG0170B. Sit to lying GG0170C. Lying to sitting on side of bed GG0170D. Sit to stand

	2321 Functional Change: Change in Mobility Score	2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients
		GG0170E. Chair/bed-to-chair transfer GG0170F. Toilet transfer GG0170G. Car transfer GG0170I. Walk 10 feet GG0170J. Walk 50 feet with two turns GG0170K. Walk 150 feet GG0170L. Walking 10 feet on uneven surfaces GG1070M. 1 step (curb) GG0170N. 4 steps GG0170O. 12 steps GG0170P. Picking up object GG0170R. Wheel 50 feet with two turns (for patients who do not walk at admission and discharge) GG0170S. Wheel 150 feet (for patients who do not walk at admission and discharge)

Appendix B2: Competing Measures (narrative format)

Comparison of 2286 and 2633

2286 Functional Change: Change in Self Care Score
2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Steward

2286 Functional Change: Change in Self Care Score

Uniform Data System for Medical Rehabilitation

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Centers for Medicare & Medicaid Services

Description

2286 Functional Change: Change in Self Care Score

Change in rasch derived values of self-care function from admission to discharge among adult patients treated at an inpatient rehabilitation facility who were discharged alive. The timeframe for the measure is 12 months. The measure includes the following 8 items: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

This measure estimates the risk-adjusted mean change in self-care score between admission and discharge for Inpatient Rehabilitation Facility (IRF) Medicare patients.

Type

2286 Functional Change: Change in Self Care Score

Outcome

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Outcome

Data Source

2286 Functional Change: Change in Self Care Score

Instrument-Based Data, Other

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Instrument-Based Data

Level

2286 Functional Change: Change in Self Care Score

Facility, Other

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Facility

Setting

2286 Functional Change: Change in Self Care Score

Inpatient/Hospital

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Post-Acute Care

Numerator Statement

2286 Functional Change: Change in Self Care Score

Average change in rasch derived self-care functional score from admission to discharge at the facility level, including items: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory. Average is calculated as: (sum of change at the patient level for all items (Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory) / total number of patients).

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

The measure does not have a simple form for the numerator and denominator. This measure estimates the risk-adjusted change in self-care score between admission and discharge among Inpatient Rehabilitation Facility (IRF) Medicare patients age 21 or older. The change in self-care score is calculated as the difference between the discharge self-care score and the admission self-care score.

Numerator Details

2286 Functional Change: Change in Self Care Score

For Inpatient Rehabilitation Facilities (IRFs) data collection currently occurs as required by the Centers for Medicare and Medicaid Services (CMS) reimbursement using the mandated payment document, the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI). Embedded in the IRF-PAI is the FIM® Instrument. The FIM® Instrument is a criterion referenced tool with 18 items that measure patient physical and cognitive functional status and patient burden of care (level of dependence/need for helper assistance). Each item is rated on a scale of 1 (most dependent) to 7 (completely independent). For the purposes of this measure, a subset of 8 FIM® items has been tested and validated which comprise the self-care measure; those items are: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory. Rasch analysis was performed on the 8 items and the difference in the rasch derived values (defined in S.2b) from admission to discharge reflect the change at the patient level. The numerator of the measure is the facility´s average change.

While the IRF-PAI is specific to inpatient rehabilitation facilities, the measure can

be used in all post-acute care venues. The FIM® instrument can be assessed in all venues of care and has been tested and validated for use in inpatient medical rehabilitation, long term acute care facilities (LTAC), skilled nursing facilities (SNF) and home health. At present, numerous LTACs and SNFs utilize the FIM® instrument (www.udsmr.org), thus the self-care measure is applicable for use in IRF, SNF, LTAC and other venues where patient functional change is anticipated.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Seven self-care activities are each scored based on a patient´s ability to complete the activity. The scores for the seven activities are summed to obtain a self-care score at the time of admission and at the time of discharge. The change in self-care is the difference between the discharge self-care score and the admission self-care score.

The 7 self-care items are:

GG0130A. Eating

GG0130B. Oral hygiene

GG0130C. Toileting hygiene

GG0130E. Shower/bathe self

GG0130F. Upper body dressing

GG0130G. Lower body dressing

GG0130H. Putting on/taking off footwear

Each patient´s ability to complete each self-care activity (item) is rated by a clinician using the following 6-level rating scale:

level 06 - Independent

level 05 - Setup or clean up assistance

level 04 - Supervision or touching assistance

level 03 - Partial/moderate assistance

level 02 - Substantial/maximal assistance

level 01 - Dependent

If the patient did not attempt the activity, the reason that the activity did not occur is reported as:

07 = Patient refused

09 = Not applicable

10 = Not attempted due to environmental limitations

88 = Not attempted due to medical condition or safety concerns.

The performance period is 12 months for reporting on CMS's IRF Compare website.

Denominator Statement

2286 Functional Change: Change in Self Care Score

Facility adjusted adjusted expected change in rasch derived values, adjusted at the

Case Mix Group level.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

The denominator is the number of Inpatient Rehabilitation Facility Medicare patient stays, except those that meet the exclusion criteria.

Denominator Details

2286 Functional Change: Change in Self Care Score

To calculate the facility adjusted expected self-care change in rasch derived values, indirect standardization is used, which weights national CMG-specific values by facility-specific CMG proportions. CMG-adjustment derives the expected value based on the case mix and severity mix of each facility. The case-mix group (CMG) classification system groups similarly impaired patients based on functional status at admission (in essence, patient severity). Patients within the same CMG are expected to have similar resource utilization needs and similar outcomes. There are three steps to classifying a patient into a CMG at admission:

1. Identify the patient's impairment group code (IGC).

2. Calculate the patient's weighted motor index score, calculated from 12 of the 13 motor FIM® items as indicated on the CMS IRF-PAI v. 20 instrument (attached).

3. Calculate the cognitive FIM® rating (as indicated on the CMS IRF_PAI v. 20 instrument) and the patient age at admission. (This step is not required for all CMGs.)

See file uploaded in S.15 for calculations.

While CMGs are only present for patients admitted to an IRF, the same procedure can be used for patients receiving care at a LTAC facility and/or a SNF, with groupings specific to those venues of care.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

The denominator is the number of Inpatient Rehabilitation Facility Medicare Part A and Medicare Advantage patient stays, except those that meet the exclusion criteria.

Exclusions

2286 Functional Change: Change in Self Care Score

National values used in the CMG-adjustment procedure will not include cases who died in the IRF or cases less than 18 years old. It is standard to exclude cases who died during rehabilitation as this is a highly atypical outcome, in addition, minors are excluded as well. The measure testing file includes further explanation regarding the exclusion criteria as well as references.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

This quality measure has six patient-level exclusion criteria:

1) Patients with incomplete stays.

Rationale: When a patient has an incomplete stay, for example, the patients leave urgently due to a medical emergency, it can be challenging to gather accurate discharge functional status data. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days.

2) Patients who are independent with all self-care activities at the time of admission.

Rationale: Patients who are independent with all the self-care items at the time of admission are assigned the highest score on all the self-care items, and thus, would not be able to show functional improvement on this same set of items at discharge.

3) Patients with the following medical conditions on admission: coma; persistent vegetative state; complete quadriplegia; locked-in syndrome; or severe anoxic brain damage, cerebral edema or compression of the brain.

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

4) Patients younger than age 21.

Rationale: There is only limited evidence published about functional outcomes for individuals with Medicare who are younger than 21.

5) Patients discharged to Hospice.

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

6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries.

Rationale: IRF-PAI data for patients not covered by the Medicare program are not submitted to the Centers for Medicare and Medicaid Services.

Facility-level quality measure exclusion: For IRFs with fewer than 20 patient stays, data for this quality measure are not publicly reported.

Exclusion Details

2286 Functional Change: Change in Self Care Score

Patient date of birth (DOB) and discharge setting are both variables collected in the IRF-PAI. Age can be calculated from DOB and patient date of admission (also collected in the IRF-PAI). In the variable discharge setting, there is a specific category for ´died´ (code: 11). Date of birth, date of admission and discharge setting (including died as a category) are also assessed in the LTAC and SNF.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

The following items are used to identify which patients are excluded from the quality measure calculations.

These data elements are included on the current version of the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), which is available

at: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/IRFPAI.html

It can be challenging to gather accurate discharge functional status data for patients who experience incomplete stays. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients discharged to a hospice; patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days.

Items used to identify these patient records:

1) Patients with incomplete stays.

Patients with a length of stay less than 3 days: We calculate length of stay using the following items on the IRF-PAI.

Item 12. Admission Date.

Item 40. Discharge Date.

Length of stay is calculated as the Discharge Date minus the Admission Date (Discharge Date - Admission Date). Patient records with a length of stay of less than 3 days are excluded.

Item 41. Patient discharged against medical advice. This item is used to identify patients discharged against medical advice.

Patient records with a response of “Yes = 1" are excluded.

Item 44C. Was the patient discharged alive? This item is used to identify patients who died during the IRF stay.

Patient records with a response of “No=0" are excluded.

44D. Patient’s discharge destination/living setting. This item is used to identify patients with an incomplete stay.

Short-term General Hospital = 02

Long-Term Care Hospital = 63

Inpatient Psychiatric Facility = 65

Critical Access Hospital = 66.

2) Patients who are independent with all self-care activities at the time of admission: Patients who are independent with all the self-care items at the time of admission are assigned the highest score on all the self-care items, and thus, would not be able to show functional improvement (i.e., a higher score)on this same set of items at discharge.

Self-care items

GG0130A. Eating = 06, and

GG0130B. Oral hygiene = 06, and

GG0130C. Toileting hygiene = 06, and

GG0130E. Shower/bathe self = 06, and

GG0130F. Upper body dressing = 06, and

GG0130G. Lower body dressing = 06, and

GG0130H. Putting on/taking off footwear = 06.

3) Patients with the following medical conditions on admission: coma; persistent vegetative state; complete quadriplegia; and locked-in syndrome; and severe anoxic brain damage, cerebral edema or compression of the brain.

The following items will be used to identify patients with these conditions:

21A. Impairment Group.

0004.1221 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia Complete, C1-C4

0004.1222 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia Complete, C5-C8

0004.2221 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C1-C4

0004.2222 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C5-C8

22. Etiologic Diagnosis.

This item is used to determine a patient´s etiologic problem that led to the condition for which the patient is receiving rehabilitation. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to identify and exclude the records of patients with these conditions:

HCC 80. Coma, Brain Compression/Anoxic Damage

ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete

ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete

ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela

ICD-10-CM. G83.5. Locked-in state

24. Comorbid Conditions.

This item is used to exclude selected comorbidities. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to identify and exclude the records of patients with these conditions:

HCC 80. Coma, Brain Compression/Anoxic Damage

ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete

ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete

ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela

ICD-10-CM. G83.5. Locked-in state

4) Patients younger than age 21.

These items are used to calculate age, and patients who are younger than 21 years of age at the time of admission are excluded.

6. Birth Date

12. Admission Date

Age is calculated as the Admission Date minus the Birth Date (Admission Date - Birth Date). Patients younger than 21 are excluded.

5) Patients discharged to hospice

44D. Patient's discharge destination/living setting.

This item is used to identify patients discharged to hospice. The following responses are used:

Hospice (home) = 50

Hospice (institutional facility) = 51

6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries.

The following items are used to identify and exclude the records of patients who are not Medicare Part A and Medicare Advantage beneficiaries:

20A. Primary Source = 99 - Not Listed AND

20B. Secondary Source = 99 - Not Listed

Risk Adjustment

2286 Functional Change: Change in Self Care Score

Stratification by risk category/subgroup

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Statistical Risk Model

Stratification

2286 Functional Change: Change in Self Care Score

While the measure can be stratified by specific impairment type (using IGC), the CMG adjustment procedure allows for the measure to be complete, accurate, and valid for all patients within the facility, excluding cases who died and excluding patient under age 18 years.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

N/A

Type Score

2286 Functional Change: Change in Self Care Score

Ratio

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Continuous variable, e.g. average

Algorithm

2286 Functional Change: Change in Self Care Score

1. Target population: Inpatient rehabilitation facility patients, skilled nursing facility short term patients, long term acute care facility patients, and home health patients.

2. Exclusions: Age less than 18 and cases who died during the episode of care.

3. Cases meeting target process: All remaining cases.

4. Outcome: Ratio of facility level average self-care change (rasch derived values) to facility CMG adjusted expected self-care change.

5. Risk adjustment: CMG adjustment using indirect standardization of the proportion of cases at the facility by CMG, and CMG specific national average of rasch derived value of self-care change.

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

We provide the detailed calculation algorithm in an attachment entitled "IRF Detailed Function QM Specifications 2633 01-07-2019" included in the Appendix.

The detailed calculation algorithm is provided to the public in the document entitled IRF Measure Calculations and Reporting User's Manual. The current version of this document is available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/IRF-Quality-Reporting-Program-Measures-Information-.html

The following are the key steps used to calculate the measure:

1) Sum the scores of the admission self-care items to create an admission self-care score for each patient, after 'activity not attempted' codes (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are recoded. (range: 7 to 42).

2) Sum the scores of the discharge self-care items to create a discharge self-care score for each patient, after 'activity not attempted' codes (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are recoded. (range: 7 to 42).

3) Identify the records of patients who meet the exclusion criteria and exclude them from analyses.

4) Calculate the difference between the admission self-care score (from step 1) and the discharge self-care score (from step 2) for each patient to create a change in self-care score for each patient.

5) Calculate an expected change in self-care score for each patient using regression coefficients from national data and each patient's admission characteristics (risk adjustors).

6) Calculate an average change in self-care score for each IRF. This is the facility-level observed change in self-care score.

7) Calculate an average expected change in self-care score for each IRF. This is the facility-level expected change in self-care score.

8) Subtract the facility-level expected change score from the facility -level observed change score to determine the difference in scores (difference value). A difference value that is 0 indicates the observed and expected scores are equal. An observed minus expected difference value that is higher than 0 (positive value)

indicates that the observed change score is greater (better) than the expected change score. An observed minus expected difference value that is less than 0 (negative value) indicates that the observed change score is lower (worse) than the expected change score.

9) Add the national average change in self-care score to each IRF's difference value (from step 8). This is the risk-adjusted mean change in self-care score.

Each patient´s ability to complete each self-care activity (item) is rated by a clinician using the following 6-level rating scale:

level 06 - Independent

level 05 - Setup or clean up assistance

level 04 - Supervision or touching assistance

level 03 - Partial/moderate assistance

level 02 - Substantial/maximal assistance

level 01 - Dependent

The 7 self-care items are:

GG0130A. Eating

GG0130B. Oral hygiene

GG0130C. Toileting hygiene

GG0130E. Shower/bathe self

GG0130F. Upper body dressing

GG0130G. Lower body dressing

GG0130H. Putting on/taking off footwear

Comparison of 2321 and 2634

2321 Functional Change: Change in Mobility Score

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Steward

2321 Functional Change: Change in Mobility Score

Uniform Data System for Medical Rehabilitation

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Centers for Medicare & Medicaid Services

Description

2321 Functional Change: Change in Mobility Score

Change in rasch derived values of mobility function from admission to discharge among adult inpatient rehabilitation facility patients aged 18 years and older who were discharged alive. The timeframe for the measure is 12 months. The measure

includes the following 4 mobility FIM® items:Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

This measure estimates the mean risk-adjusted mean change in mobility score between admission and discharge for Inpatient Rehabilitation Facility (IRF) Medicare Part A and Medicare Advantage patients.

Type

2321 Functional Change: Change in Mobility Score

Outcome

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Outcome

Data Source

2321 Functional Change: Change in Mobility Score

Instrument-Based Data, Other

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Instrument-Based Data

Level

2321 Functional Change: Change in Mobility Score

Facility, Other

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Facility

Setting

2321 Functional Change: Change in Mobility Score

Inpatient/Hospital, Post-Acute Care

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Post-Acute Care

Numerator Statement

2321 Functional Change: Change in Mobility Score

Average change in rasch derived mobility function score from admission to discharge at the facility level. Includes the following items: Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs. Average is calculated as: (sum of change at the patient level/total number of patients). Patient less than 18 years of age at admission to the facility or patients who died within the facility are excluded.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

The measure does not have a simple form for the numerator and denominator. This measure estimates the risk-adjusted change in mobility score between admission and discharge among Inpatient Rehabilitation Facility (IRF) Medicare Part A and Medicare Advantage patients age 21 and older. The change in mobility score is calculated as the difference between the discharge mobility score and the admission mobility score.

Numerator Details

2321 Functional Change: Change in Mobility Score

For Inpatient Rehabilitation Facilities (IRFs) data collection is presently required for payment reimbursement by the Centers for Medicare and Medicaid Services (CMS) using the mandated Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI). Embedded in the IRF-PAI is the FIM® Instrument. The FIM® Instrument is a criterion referenced tool with 18 items that measures patient physical and cognitive function, need for helper assistance, burden of care/level of dependence. Each item is rated on a scale of 1 (most dependent) to 7 (completely independent). For the purposes of this measure, a subset of 4 FIM® items has been tested and validated as the Change in Mobility measure; the items are: Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs. Rasch analysis was performed on the items and the difference in the rasch derived values (defined in S.2b) from admission to discharge reflect the change at the patient level. The numerator of the measure is the average change in mobility score at the facility level.

While the IRF-PAI is specific to inpatient rehabilitation facilities, the change in mobility measure can be used in all post-acute care venues. The FIM® instrument is routinely used for patient functional assessment in all venues of care and has been tested and validated for use in IRFs, skilled nursing facilities (SNFs) and long term acute care facilities (LTAC) (www.udsmr.org), therefore this measure is not specific for inpatient medical rehabilitation use only.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Seventeen mobility activities are each scored based on a patient´s ability to complete the activity. The scores for the activities are summed to obtain a mobility score at the time of admission and at the time of discharge. The change in mobility is the difference between the discharge mobility score and the admission mobility score.

The mobility items are:

GG0170A. Roll left and right

GG0170B. Sit to lying

GG0170C. Lying to sitting on side of bed

GG0170D. Sit to stand

GG0170E. Chair/bed-to-chair transfer

GG0170F. Toilet transfer

GG0170G. Car transfer

GG0170I. Walk 10 feet

GG0170J. Walk 50 feet with two turns

GG0170K. Walk 150 feet

GG0170L. Walking 10 feet on uneven surfaces

GG1070M. 1 step (curb)

GG0170N. 4 steps

GG0170O. 12 steps

GG0170P. Picking up object

GG0170R. Wheel 50 feet with two turns (for patients who do not walk at admission and discharge)

GG0170S. Wheel 150 feet (for patients who do not walk at admission and discharge)

Each patient´s ability to complete each mobility activity (item) is rated by a clinician using the following 6-level rating scale:

level 06 - Independent

level 05 - Setup or clean up assistance

level 04 - Supervision or touching assistance

level 03 - Partial/moderate assistance

level 02 - Substantial/maximal assistance

level 01 - Dependent

If the patient did not attempt the activity, the reason that activity did not occur is reported as:

07 = Patient refused

09 = Not applicable

10 = Not attempted due to environmental limitations

88 = Not attempted due to medical condition or safety concerns.

The performance period is 12 months for reporting on CMS's IRF Compare website.

Denominator Statement

2321 Functional Change: Change in Mobility Score

Facility adjusted expected change in rasch derived mobility values, adjusted at the Case Mix Group (CMG) level.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

The denominator is the number of Inpatient Rehabilitation Facility Medicare Part A and Medicare Advantage patient stays, except those that meet the exclusion criteria.

Denominator Details

2321 Functional Change: Change in Mobility Score

To calculate the facility adjusted expected change in rasch derived mobility values, indirect standardization was used, which weights national CMG-specific values by facility-specific CMG proportions. CMG-adjustment derives the expected value based on the case mix and severity mix of each facility. The case-mix group (CMG) classification system groups similarly impaired patients based on functional status at admission, in essence, patient severity. Patients within the same CMG are expected to have similar resource utilization needs and similar functional outcomes. There are three steps to classifying a patient into a CMG at admission:

1. Identify the patient's impairment group code (IGC).

2. Calculate the patient's weighted motor index score, calculated from 12 of the 13 motor FIM® items.

3. Calculate the cognitive FIM® rating and the patient´s age at admission. (This step is not required for all CMGs.)

See file uploaded in S.2b for calculations or ´CMG Version 3.00 [ZIP, 9.02mb]´ at the following link for more details:

https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/CMG.html

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

The denominator is the number of Inpatient Rehabilitation Facility Medicare Part A and Medicare Advantage patient stays, except those that meet the exclusion criteria.

Exclusions

2321 Functional Change: Change in Mobility Score

National values used in the CMG adjustment procedure will not include cases who died in the IRF or patients less than 18 years of age at admission. Cases who died during rehabilitation are not typical patients and are routinely omitted from reports and published research on rehabilitation outcomes. Further details and references related to the exclusion criteria can be found in the Measure Testing form.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

This quality measure has six patient-level exclusion criteria:

1) Patients with incomplete stays.

Rationale: When a patient has an incomplete stay, for example, the patients leave urgently due to a medical emergency, it can be challenging to gather accurate discharge functional status data. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care

Hospital); patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days.

2) Patients who are independent with all mobility activities at the time of admission.

Rationale: Patients who are independent with all the mobility items (with the exception of the wheelchair items GG0170R and GG0170S) at the time of admission are assigned the highest score on all the mobility items, and thus, would not be able to show functional improvement on this same set of items at discharge.

3) Patients with the following medical conditions on admission: coma, persistent vegetative state; complete quadriplegia; locked-in syndrome or severe anoxic brain damage, cerebral edema or compression of brain.

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

4) Patients younger than age 21.

Rationale: There is only limited evidence published about functional outcomes for individuals with Medicare who are younger than 21.

5) Patients discharged to hospice.

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

6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries.

Rationale: IRF-PAI data for patients not covered by the Medicare program are not submitted to the Centers for Medicare and Medicaid Services.

Facility-level quality measure exclusion: For IRFs with fewer than 20 patient stays, data for this quality measure are not publicly reported.

Exclusion Details

2321 Functional Change: Change in Mobility Score

Patient date of birth (DOB), date of admission and discharge setting variables are collected in the IRF-PAI. Age can be calculated from DOB and admission date. The variable discharge setting includes a category for ´died´ which is indicated as a code of ´11´. Patient date of birth, admission date and discharge setting are also documented in SNFs and LTAC facilities.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

The following items are used to identify which patients are excluded from the quality measure calculations.

These data elements are included on the current version of the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), which is available at: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/IRFPAI.html

It can be challenging to gather accurate discharge functional status data for patients who experience incomplete stays. Patients with incomplete stays include

patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital); patients discharged to a hospice; patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days.

Items used to identify these patient records:

1) Patients with incomplete stays.

Patients with a length of stay less than 3 days: We calculate length of stay using the following items on the IRF-PAI.

Length of stay is calculated as the Discharge Date minus the Admission Date (Discharge Date - Admission Date). Patient records with a length of stay of less than 3 days are excluded.

Item 12. Admission Date.

Item 40. Discharge Date.

Item 41. Patient discharged against medical advice. This item is used to identify patients discharged against medical advice.

Patient records with a response of "Yes = 1" are excluded.

Item 44C. Was the patient discharged alive? This item is used to identify patients who died during the IRF stay.

Patient records with a response of "No = 0" are excluded.

44D. Patient's discharge destination/living setting.

This item is used to identify an incomplete stay. Specifically, the following responses will be used to identity patients with incomplete stays:

Short-term General Hospital = 02

Long-Term Care Hospital = 63

Inpatient Psychiatric Facility = 65

Critical Access Hospital = 66.

2) Patients who are independent with all mobility activities at the time of admission.

Patients who are independent with all the mobility items at the time of admission are assigned the highest score on all the mobility items, thus, would not be able to show functional improvement (i.e., a higher score) on this same set of items at discharge. The following items and scores are used to identify and exclude patient records:

Mobility items

GG0170A. Roll left and right = 06, and

GG0170B. Sit to lying = 06, and

GG0170C. Lying to sitting on side of bed = 06, and

GG0170D. Sit to stand = 06, and

GG0170E. Chair/bed-to-chair transfer = 06, and

GG0170F. Toilet transfer = 06, and

GG0170G. Car transfer = 06, and

GG0170I. Walk 10 feet = 06, and

GG0170J. Walk 50 feet with two turns = 06, and

GG0170K. Walk 150 feet = 06, and

GG0170L. Walking 10 feet on uneven surfaces = 06, and

GG0170M. 1 step (curb) = 06, and

GG0170N. 4 steps = 06, and

GG0170O. 12 steps = 06, and

GG0170P. Picking up object = 06.

3) Patients with the following medical conditions on admission: coma; persistent vegetative state; complete quadriplegia; locked-in syndrome; and severe anoxic brain damage, cerebral edema or compression of the brain.

The following items will be used to identify patients with these conditions:

21A. Impairment Group.

0004.1221 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia Complete, C1-C4

0004.1222 - Spinal Cord Dysfunction, Non-Traumatic: Quadriplegia Complete, C5-C8

0004.2221 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C1-C4

0004.2222 - Spinal Cord Dysfunction, Traumatic: Quadriplegia Complete, C5-C8

22. Etiologic Diagnosis.

This item is used to determine a patient´s etiologic problem that led to the condition for which the patient is receiving rehabilitation. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to identify and exclude records of patients with these conditions:

HCC 80. Coma, Brain Compression/Anoxic Damage

ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete

ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete

ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela

ICD-10-CM. G83.5. Locked-in state

24. Comorbid Conditions.

This item is used to exclude selected comorbidities. The following Hierarchical Condition Categories (HCCs) and International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes will be used to exclude records of patients with these conditions:

HCC 80. Coma, Brain Compression/Anoxic Damage

ICD-10-CM. G82.51 Quadriplegia, C1-C4 complete

ICD-10-CM. G82.53 Quadriplegia, C5-C7 complete

ICD-10-CM. S14.11xx Quadriplegia, Complete lesion at Cx(1-8) level of cervical spinal cord, initial encounter or subsequent encounter, or sequela

ICD-10-CM. G83.5. Locked-in state

4) Patients younger than age 21. These items are used to calculate age, and patients who are younger than 21 years of age at the time of admission are excluded.

6. Birth Date

12. Admission Date

Age is calculated as the Admission Date minus the Birth Date (Admission Date - Birth Date). Patients younger than 21 are excluded.

5) Patients discharged to hospice.

44D. Patient's discharge destination/living setting.

This item is used to identify patients discharged to hospice. The following responses are used:

Hospice (home) = 50

Hospice (institutional facility) = 51

6) Patients who are not Medicare Part A and Medicare Advantage beneficiaries

20A. Primary Source = 99 - Not Listed AND

20B. Secondary Source = 99 - Not Listed

Risk Adjustment

2321 Functional Change: Change in Mobility Score

Stratification by risk category/subgroup

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Statistical Risk Model

Stratification

2321 Functional Change: Change in Mobility Score

While the measure can be stratified by specific impairment type (IGC), the CMG adjustment procedure allows for the measure to be complete, accurate, and valid for all patients within the facility, excluding died cases and ages less than 18.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

N/A

Type Score

2321 Functional Change: Change in Mobility Score

Ratio

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Continuous variable, e.g. average

Algorithm

2321 Functional Change: Change in Mobility Score

1. Target population: patients receiving care at an inpatient medical rehabilitation facility, a skilled nursing facility, or a long term acute care facility.

2. Exclusions: Age less than 18 years and patients who died during the episode of care.

3. Cases meeting target process: All remaining cases.

4. Outcome: Ratio of facility level average mobility change (rasch derived values) to facility CMG adjusted expected mobility change.

5. Risk adjustment: CMG adjustment using indirect standardization of the proportion of cases at the facility by CMG, and CMG specific national average of rasch derived value of mobility change.

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

We provide the detailed calculation algorithm in an attachment entitled "IRF Detailed Function QM Specifications 2634 01-07-2019" included in the Appendix.

The detailed calculation algorithm is provided to the public in the document entitled IRF Measure Calculations and Reporting User's Manual. The current version of this document is available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/IRF-Quality-Reporting/IRF-Quality-Reporting-Program-Measures-Information-.html

The following are key steps used to calculate the measure:

1) Sum the scores of the admission mobility items to create an admission mobility score for each patient, after 'activity not attempted' codes (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are recoded, and for patients who do not walk on admission and discharge, walking items have been recoded to use wheelchair mobility item codes. (range: 15 to 90).

2) Sum the scores of the discharge mobility items to create a discharge mobility score for each patient, after 'activity not attempted' values (07. Patient refused, 09. Not applicable, 10. Not attempted due to environmental limitations, and 88. Not attempted due to medical condition or safety concerns), skip codes ('^') and missing data ('-') are recoded. As described in step 1, for patients who do not walk on admission and discharge, use wheelchair mobility item codes instead of walking codes. (range: 15 to 90).

3) Identify the records of patients who meet the exclusion criteria and exclude them from analyses.

4) Calculate the difference between the admission mobility score (from step 1) and the discharge mobility score (from step 2) for each patient to create a change in mobility score for each patient.

5) Calculate an expected change in mobility score for each patient using regression coefficients from national data and each patient's admission characteristics (risk adjustors).

6) Calculate an average observed change in mobility score for each IRF (using the patient data calculated in step 4). This is the facility-level observed change in mobility score.

7) Calculate an average expected change in mobility score for each IRF (using the patient data from step 5). This is the facility-level expected change in mobility score.

8) Subtract the facility-level expected change score from the facility-level observed change score to determine the difference in scores (difference value). A difference value that is 0 indicates the observed and expected scores are equal. An observed minus expected difference value that is higher than 0 (positive) indicates that the observed change score is greater (better) than the expected change score. An observed minus expected difference value that is less than 0 (negative) indicates that the observed change score is lower (worse) than the expected change score.

9) Add the national average change in mobility score to each IRF's difference value (from step 8). This is the risk-adjusted mean change in mobility score.

Each patient´s ability to complete each mobility activity (item) is rated by a clinician using the following 6-level rating scale:

level 06 - Independent

level 05 - Setup or clean up assistance

level 04 - Supervision or touching assistance

level 03 - Partial/moderate assistance

level 02 - Substantial/maximal assistance

level 01 - Dependent

The mobility items are:

GG0170A. Roll left and right

GG0170B. Sit to lying

GG0170C. Lying to sitting on side of bed

GG0170D. Sit to stand

GG0170E. Chair/bed-to-chair transfer

GG0170F. Toilet transfer

GG0170G. Car transfer

GG0170I. Walk 10 feet

GG0170J. Walk 50 feet with two turns

GG0170K. Walk 150 feet

GG0170L. Walking 10 feet on uneven surfaces

GG1070M. 1 step (curb)

GG0170N. 4 steps

GG0170O. 12 steps

GG0170P. Picking up object

GG0170R. Wheel 50 feet with two turns (for patients who do not walk at admission and discharge)

GG0170S. Wheel 150 feet (for patients who do not walk at admission and discharge)

Appendix C: Comments Received on Competing Measures

Comment submitted by Peg Graham on 2321: Functional Change: Change in Mobility Score and 2286: Functional Change: Change in Self Care Score:

As a family caregiver, I have been following the conversation re self-care/mobility scores across the care continuum, including discharge to community. In the Fall 2017 report of the Patient Experience and Function Standing Committee, there appeared to be uncertainty about the merits of Section GG vs FIMS. I've searched the Fall 2018 report and have had a hard time discerning whether of not this issue has been settled. In the event that the Standing Committee is still accepting comments on this issue, I urge that Section GG be selected. The Section GG 6pt scale clearly communicates the level to which a patient relies on personal assistance in a manner that the patient, clinician and family member can understand. Particularly in discharges to home, the family needs to appreciate the degree to which their loved one will be depending on their presence to perform self-care tasks and mobility/transfers.

Please note: Study examined how similar summary scores of physical functioning using the Functional Independence Measure (FIM) can represent different patient clinical profiles. Data were analyzed for 765,441 Medicare fee-for-service beneficiaries discharged from inpatient rehabilitation. Patients' scores on items of the FIM were used to quantify their level of independence on both self-care and mobility domains. Patients requiring "no physical assistance" at discharge from inpatient rehabilitation were identified by using a rule and score-based approach. In patients with FIM self-care and mobility summary scores suggesting no physical assistance needed, the study found that physical assistance was in fact needed frequently in bathroom-related activities (e.g., continence, toilet and tub transfers, hygiene, clothes management) and with stairs. It was not uncommon for actual performance to be lower than what may be suggested by a summary score of those domains. The authors conclude that further research is needed to create clinically meaningful descriptions of summary scores from combined performances on individual items of physical functioning. Citation: Fisher, Steve R., Middleton, Addie, Graham, James E., Ottenbacher, Kenneth J.. (2018). Same but different: FIM summary scores may mask variability in physical functioning profiles. Archives of Physical Medicine and Rehabilitation, 99(8), Pgs. 1479-1482, 1482.e1. Retrieved 12/6/2018, from REHABDATA database.

Appendix D: 2015 Project Report

Below are the Appendix A sections for each of the competing measures from the 2015 report.

2286 Functional Change: Change in Self Care Score

Submission

Description: Change in rasch derived values of self-care function from admission to discharge among adult patients treated at an inpatient rehabilitation facility who were discharged alive. The timeframe for the measure is 12 months. The measure includes the following 8 items: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory.

Numerator Statement: Average change in rasch derived self-care functional score from admission to discharge at the facility level, including items: Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory. Average is calculated as: (sum of change at the patient level for all items (Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, and Memory) / total number of patients).

Denominator Statement: Facility adjusted adjusted expected change in rasch derived values, adjusted at the Case Mix Group level.

Exclusions: National values used in the CMG-adjustment procedure will not include cases who died in the IRF (or other venue) or cases less than 18 years old. Cases who died during rehabilitation are not typical patients and are typically omitted in the literature when looking at rehabilitation outcomes. In addition, the FIM instrument is meant for an adult population (Ottenbacher et al. 1996).

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Home Health, Post Acute/Long Term Care Facility : Inpatient Rehabilitation Facility, Post Acute/Long Term Care Facility : Long Term Acute Care Hospital, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility

Type of Measure: Outcome

Data Source: Electronic Clinical Data : Electronic Health Record, Other

Measure Steward: Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities, Inc. and its successor in interest, UDSMR, LLC.

STANDING COMMITTEE MEETING [01/21/2015-01/22/2015]

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap, 1c. High Impact)

1a. Evidence: **Y-17**; **N-0**; 1b. Performance Gap: **H-4**; **M-8**; **L-0**; **I-5;** 1c. High Priority: **H-9**; **M-8**; **L-0**; **I-0**

Rationale:

- This is one of a suite of measures derived from the FIM. This measure, 2286, calculates and reports a change in self-care score; measure 2321 reports a change in mobility

score, and together they comprise measure 2287, which calculates a change in motor score. The developer explained they are proposing three measures because different aspects of the measures (self-care indicators vs. mobility indicators) could differ in importance based on the setting and the patient's prognosis or condition.
- The Committee inquired about the lack of information on disparities in measure performance; the developer indicated the data is available; however, due to the wealth of information they have, they were unsure how much and what data to submit. They agreed to provide additional information, specifically on age, race and payer source, during the public comment period.
- The Committee requested clarification on the measure timing requirements of one year; the developer responded that the assessments occur at admission and discharge, regardless of the length of stay. That the one-year period was a mechanism to assess facility performance for patients who have both the admission and discharge scores and then compare against benchmarks.
- The developer also explained that the FIM allows assessment of both function and burden of care. Burden of care refers to how much time a patient would require from a helper, another person, or one-on-one, if living within a community setting.
- The measure is not restricted to Medicare-only but can include patients starting at 18 years of age.
- There was discussion about the appropriate setting of care for measure implementation, and while the developers indicated it can be used across various settings, the data provided was only for IRF's. Thus the Committee was instructed to evaluate and vote based on the data and specification submitted which was specific to IRFs.
- The Committee clarified that expression and memory are components of the self-care metric.
- The Committee proposed that the votes for measure 2286 be carried over to measures 2287 and 2321.

---

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **H-6**; **M-6**; **L-1**; **I-4** 2b. Validity: **H-4**; **M-9**; **L-0**; **I-4**

<u>Rationale</u>:

- It was noted that these are clinician-derived scores which require fairly rigorous training of appropriate clinicians to ensure reliability.
- The Committee clarified that sufficient evidence was provided for reliability at the patient level, but not at the agency level. The developer confirmed this interpretation and indicated the availability of additional information to be supplied during the public comment period.
- The Committee inquired if the testing results were based on raw scores versus the Rasch-transformed scores. It was noted that the impact of change could differ based on the use of the raw scores. The developer indicated that by converting to Rasch scores, it helped to mitigate drastic differences. The data provided was all Rasch-transformed, and they are able to provide the raw data detail as well.

- The Committee requested clarification on the risk adjustment methodology. The developer starts by classifying patients into an impairment group and then calculates the patient score. They then proceed to look at facility case-mix; then make a final adjustment to have a facility adjusted score, in addition to the patient adjusted score. By adjusting at both levels, the results are comparable between facilities and between patients.
- The Committee clarified their request for data and asked for the Interclass Correlation Coefficients, as well as mean square fit statistics.
- The Committee asked for additional information regarding the testing of 4 items correlated with the overall FIM since the result was .60. The developer indicated they specifically looked at the 4 items and assessed how they predict the patient's full 18-item FIM score and felt the results were reasonable. It was confirmed that they were looking at validity and the proportion of variance that was accounted for in those 4 items. The Committee suggested that over time, the measure may be better off with the 2-subscales as more valid overall.

---

4. Feasibility: H-3; M-11; L-3; I-0

(4a. Clinical data generated during care delivery; 4b. Electronic sources; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)

Rationale:

- As discussed under reliability, the Committee raised the importance of proper training for clinicians using this tool. The developer indicated there are training modules available and variations in training systems (i.e., train the trainer).
- There was concern raised about feasibility in settings outside of the IRF; and although the developer indicates potential for wider spread use, the measure as submitted for Committee consideration is for IRFs only.

---

3. Use and Usability: H-6; M-9; L-0; I-2

(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting/Accountability and 3b. Quality Improvement)

Rationale:

- The Committee requested clarification on the availability of data for accountability and benchmarking. The developer confirmed that the benchmarking piece is not publicly available.
- It was noted that CMS conducts a significant amount of oversight on these facilities.

---

5. Related and Competing Measures

- The Committee considered this measure to potentially compete with 2633: IRF Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients (CMS) and was asked to vote to determine whether these measures are directly competing and select the best in class measure. While the Committee agreed that these measures are competing, they did not achieve consensus on whether one measure was

superior. When measures 2286 and 2633 moved forward, the CSAC voted to recommend 2286, but not 2633, as the measure was deemed as competing with 2286. The Board of Directors reviewed the recommendations of the CSAC and the rationale for non-approval of 2633. The Board provided greater policy context, including the importance of the IMPACT Act of 2014 and the need for aligned measures that can be used to assess care across settings. The Board therefore directed NQF staff to return the competing IRF measures (2286 and 2633) back to the CSAC for further consideration. In addition, the Board expressed concerns regarding measures derived from proprietary versus non-proprietary instruments, and the desirability of having measures that help assess quality improvement from the patient's perspective as he/she moves among multiple sites of care. In their reconsideration vote, 92% of the CSAC voted to approve endorsement for both measures with conditions for specific update requirements. The Committee also considered this measure to be related to 2635: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients (CMS) and 2613: CARE: Improvement in Self Care (AHCA); however, there were no recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-15**; **N-2**

6. Public and Member Comment: March 2, 2015- March 31, 2015

Comments received:

- Two sets of comments suggested that 2286, 2287, and 2321 be harmonized. As this decision is up to the developer, these comments were forwarded for their response.

Developer response:

We appreciate the endorsement. We agree that a composite measure is important. To that end, we have submitted a composite measure 2287: Functional Change: Change in Motor Score. This will allow for quality improvement in all levels of function being measured. However, we feel that leaving this as a separate measure offers greater refinement in assessing patient change relating to the construct measured. For instance, consider a patient admitted to a facility and upon admission is rated at the lowest functional levels for each item within a measure, upon discharge, the self-care items improved greatly however the mobility items did not change from the admission rating (perhaps the patient had not walked independently for many years prior to onset of recent condition under treatment), as a composite score, functional gain would be evident from admission to discharge, but it would not show the domain specific changes (exceptional progress in self-care, which was likely the focus of rehabilitation). We believe the option of serving as a 'stand alone measure' may have interest and great utility to clinicians and since the motor measure is a combination of the self-care and mobility, the flexibility in options exist for clinical use.

7. Review 1: Consensus Standards Approval Committee (CSAC) Vote June 29, 2015: Y-14; N-1; A-0

Review 2: CSAC Vote September 17, 2015: Y-12; N-1; A-0

CSAC Decision: Approved for endorsement with conditions for updates – Final Decision made September 17, 2015

8. Review 1: Board of Directors Review: No (July 22, 2015)

Board Decision: The Board decided to send Measure 2286 back to the CSAC for further consideration.

Review 2: Board of Directors Review: Yes (November 4, 2015)

Board Decision: Ratified for endorsement– Final Decision made November 4, 2015

9. Appeals

2321 Functional Change: Change in Mobility Score

Submission

Description: Change in rasch derived values of mobility function from admission to discharge among adult inpatient rehabilitation facility patients aged 18 years and older who were discharged alive. The timeframe for the measure is 12 months. The measure includes the following 4 mobility FIM® items:Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs.

Numerator Statement: Average change in rasch derived mobility functional score from admission to discharge at the facility level. Includes the following FIM items: Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs. Average is calculated as (sum of change at the patient level/total number of patients). Cases aged less than 18 years at admission to the facility or patients who died within the facility are excluded.

Denominator Statement: Facility adjusted adjusted expected change in rasch derived values, adjusted at the Case Mix Group level.

Exclusions: National values used in the CMG-adjustment procedure will not include cases who died in the IRF (or other venue) or cases less than 18 years old. Cases who died during rehabilitation are not typical patients and are typically omitted in the literature when looking at rehabilitation outcomes. In addition, the FIM instrument is meant for an adult population (Ottenbacher et al. 1996).

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Post Acute/Long Term Care Facility : Inpatient Rehabilitation Facility, Post Acute/Long Term Care Facility : Long Term Acute Care Hospital, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility

Type of Measure: Outcome

Data Source: Electronic Clinical Data : Electronic Health Record

Measure Steward: Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities, Inc. and its successor in interest, UDSMR, LLC.

STANDING COMMITTEE MEETING [01/21/2015-01/22/2015]

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap, 1c. High Impact)

1a. Evidence: **Y-17**; **N-0**; 1b. Performance Gap: **H-4**; **M-8**; **L-0**; **I-5;** 1c. High Priority: **H-9**; **M-8**; **L-0**; **I-0**

Rationale:

- This is one of a suite of measures derived from the FIM Measure 2286 calculates and reports a change in self-care score; this measure, 2321, reports a change in mobility score, and together they comprise measure 2287 which calculates a change in motor score. The developer indicated it was important of the committee to understand this and why they are proposing three measures. Different aspects of the measure (self-care indicators vs. mobility indicators) could differ in importance based on the setting and the patient prognosis or condition.
- The Committee inquired as to the lack of information on disparities in measure performance; the developer indicated the data is available, however, due to the wealth of information they have, they were unsure how much and what data to submit. They agreed to provide additional information, specifically on age, race and payer source, during the public comment period.
- The Committee requested clarification on the measure timing requirements of one year; the developer responded that the assessments occur at admission and discharge, regardless of the length of stay. That the one-year period was a mechanism to assess facility performance for patients who have both the admission and discharge scores and then compare against benchmarks.
- The developer also explained that the FIM allows assessment of both function and burden of care. Burden of care refers to how much time a patient would require from a helper, another person, or one-on-one if living within a community setting.
- The measure is not restricted to Medicare-only but can include patients starting at 18 years of age.
- There was discussion about the appropriate setting of care for measure implementation, and while the developers indicated it can be used across various settings, the data provided was only for IRF's. Thus the Committee was instructed to evaluate and vote based on the data and specification submitted which was specific to IRFs.
- The Committee clarified that expression and memory are components of the self-care metric.
- The Committee proposed that the votes for measure 2286 be carried over to measures 2287 and 2321.

2. Scientific Acceptability of Measure Properties: The measure meets the Scientific Acceptability criteria

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **H-6**; **M-6**; **L-1**; **I-4** 2b. Validity: **H-4**; **M-9**; **L-0**; **I-4**

Rationale:

- It was noted that these are clinician derived scores which require fairly rigorous training of appropriate clinicians to ensure reliability.
- The Committee clarified that sufficient evidence was provided for reliability at the patient level, but the agency level data included a beta binomial model and the interclass correlation coefficients look like a measure level mean variance. These rates were used to estimate rates as opposed to the composite score which is what would be sued to evaluate performance of the agencies. Thus, the interclass correlations are at the measure level versus the facility level. The developer confirmed this interpretation and indicated the availability of additional information to be supplied during the Public Comment period.
- The Committee inquired if the testing results were based on raw scores versus the Rasch-transformed scores. It was noted that the impact of change could differ based on the use of the raw scores. The developer indicated that by converting to Rasch scores, it helped to mitigate drastic differences. The data provided was all Rasch-transformed, and they are able to provide the raw data detail as well.
- The Committee requested clarification on the risk adjustment methodology. The developer starts by classifying patients into an impairment group and then calculates the patient score. They then proceed to look at facility case-mix; then make a final adjustment to have a facility adjusted score, in addition to the patient adjusted score. By adjusting at both levels, the results are comparable between facilities and between patients.
- The Committee clarified their request for data and asked for the Interclass Correlation Coefficients, as well as mean square fit statistics.

---

4. Feasibility: H-3; M-11; L-3; I-0

(4a. Clinical data generated during care delivery; 4b. Electronic sources; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)

Rationale:

- As discussed under reliability, the Committee raised the importance of proper training for clinicians using this tool. The developer indicated there are training modules available and variations in training systems (i.e., train the trainer)
- There was concern raised about feasibility in settings outside of the IRF; and although the developer indicates potential for wider spread use, the measure as submitted for Committee consideration is for IRFs only.

---

3. Use and Usability: H-6; M-9; L-0; I-2

(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting/Accountability and 3b. Quality Improvement)

Rationale:

- The Committee requested clarification on the availability of data for accountability and benchmarking. The developer confirmed that the benchmarking piece is not publicly available.

5. Related and Competing Measures

- The Committee considered this measure to be related to 2612: CARE: Improvement in Mobility (AHCA), 2632: Long-Term Care Hospital (LTCH) Functional Outcome Measure: Change in Mobility Among Patients Requiring Ventilator Support (CMS), and 2636: IRF Functional Outcome Measure: Discharge Mobility Score for Medical Rehabilitation Patients (CMS). These measures have the same focus area (mobility) but are specified for different types of target populations. The Committee agreed that there was a need for all of the aforementioned measures, but made no recommendations for harmonization.
- The Committee considered this measure to potentially compete with 2634: IRF Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients (CMS) and was asked to vote to determine whether these measures are directly competing and select the best in class measure.
- While the Committee agreed that these measures are competing, they did not achieve consensus on whether one measure was superior. When measures 2321 and 2634 moved forward, the CSAC voted to recommend 2321, but not 2634, as the measure was deemed as competing with 2321. The Board of Directors reviewed the recommendations of the CSAC and the rationale for non-approval of 2634. The Board provided greater policy context, including the importance of the IMPACT Act of 2014 and the need for aligned measures that can be used to assess care across settings. The Board therefore directed NQF staff to return the competing IRF measures (2321 and 2634) back to the CSAC for further consideration. In addition, the Board expressed concerns regarding measures derived from proprietary versus non-proprietary instruments, and the desirability of having measures that help assess quality improvement from the patient's perspective as he/she moves among multiple sites of care. In their reconsideration vote, 92% of the CSAC voted to approve endorsement for both measures with conditions for specific update requirements.

Standing Committee Recommendation for Endorsement: **Y-15**; **N-2**

6. Public and Member Comment: March 2, 2015- March 31, 2015

Comments received:

- Two sets of comments suggested that 2286, 2287, and 2321 be harmonized. As this decision is up to the developer, these comments were forwarded on for their response.

Developer response:

- We appreciate the endorsement. We agree that a composite measure is important. To that end, we have submitted a composite measure 2287: Functional Change: Change in Motor Score. This will allow for quality improvement in all levels of function being measured. However, we feel that leaving this as a separate measure offers greater refinement in assessing patient change relating to the construct measured. For instance, consider a patient admitted to a facility and upon admission is rated at the lowest functional levels for each item within a measure, upon discharge, the self-care items improved greatly however the mobility items did not change from the admission rating (perhaps the patient had not walked independently for many years prior to onset of

recent condition under treatment), as a composite score, functional gain would be evident from admission to discharge, but it would not show the domain specific changes (exceptional progress in self-care, which was likely the focus of rehabilitation). We believe the option of serving as a 'stand alone measure' may have interest and great utility to clinicians and since the motor measure is a combination of the self-care and mobility, the flexibility in options exist for clinical use.

---

7. Review 1: Consensus Standards Approval Committee (CSAC) Vote June 29, 2015: Y-14; N-1; A-0

Review 2: CSAC Vote September 17, 2015: Y-12; N-1; A-0

---

8. Review 1: Board of Directors Review: No (July 22, 2015)

Board Decision: The Board decided to send Measure 2321 back to the CSAC for further consideration.

Review 2: Board of Directors Review: Yes (November 4, 2015)

Board Decision: Ratified for endorsement, with conditions for updates– Final Decision made November 4, 2015

2633 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients

Submission

Description: This measure estimates the risk-adjusted mean change in self-care score between admission and discharge for Inpatient Rehabilitation Facility (IRF) Medicare patients.

Numerator Statement: The measure does not have a simple form for the numerator and denominator. This measure estimates the risk-adjusted change in self-care score between admission and discharge among Inpatient Rehabilitation Facility (IRF) Medicare patients age 21 or older. The change in self-care score is calculated as the difference between the discharge self-care score and the admission self-care score.

Denominator Statement: Inpatient Rehabilitation Facility patients included in this measure are at least 21 years of age, Medicare beneficiaries, are not independent on all of the self-care activities at the time of admission, and have complete stays.

Exclusions: This quality measure has 6 exclusion criteria:

1) Patients with incomplete stays.

Rationale: It can be challenging to gather accurate discharge functional status data for patients who experience incomplete stays. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital), because of a medical emergency; patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; patients discharged directly to another IRF and patients with a length of stay less than 3 days.

2) Patients who are independent with all self-care activities at the time of admission.

Rationale: Patients who are independent with all the self-care items at the time of admission are assigned the highest score on all the self-care items, and thus, would not be able to show functional improvement on this same set of items at discharge.

3) Patients with the following medical conditions: coma; persistent vegetative state; complete tetraplegia; locked-in syndrome; or severe anoxic brain damage, cerebral edema or compression of brain.

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

4) Patients younger than age 21.

Rationale: There is only limited evidence published about functional outcomes for children.

5) Patients discharged to hospice.

Rationale: Patient goals may change during the IRF stay.

6) Patients who are not Medicare beneficiaries.

Patients not covered by the Medicare program.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Post Acute/Long Term Care Facility : Inpatient Rehabilitation Facility

Type of Measure: Outcome

Data Source: Electronic Clinical Data

Measure Steward: Centers for Medicare & Medicaid Services (CMS)

STANDING COMMITTEE MEETING [01/21/2015-01/22/2015]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

(1a. Evidence, 1b. Performance Gap, 1c. High Impact)

1a. Evidence: **Y-15**; **N-1**; 1b. Performance Gap: **H-2**; **M-12**; **L-1**; **I-1;** 1c. High Priority: **H-9**; **M-6**; **L-1**; **I-0**

<u>Rationale</u>:

- The Committee noted that the measure is proposed for use for Medicare only, and felt that this limits the use of the measure and potentially introduces duplication of efforts if using multiple tools for differing payer populations.
- The Committee requested clarification on the intent of the measure and if it was a reflection of the care in the IRF or how the patient was prepare for integration back into the community. Specifically, they wanted to know if there is a connection between how a patient is doing at discharge and how they will do in the community. The developer indicated that information was provided in the supplemental information specific to the evidence behind the measure. CMS further explained this is another attempt to standardize measurement and allow tracking of patients as they traverse the care continuum and between settings. The measures allow the comparison of uniform assessment data, whether it's self-care or mobility.
- The Committee asked for the reasoning behind the proposal of four measures using essentially the same data. The developer indicated that when testing understanding of the measures with consumers, they were led to develop both a change score concept

for use by facilities and then the percentage of patients that achieve a certain status to improve consumer understanding. They would have provided both in the same measure if the NQF submissions allowed. There was a suggestion that these two pairs of measures be considered "paired" measure to promote their use together. A member from the rehabilitation community indicated he would find the information provided from both levels of measurement useful. Internally they can be used for the facility for quality improvement and externally for use with consumers.

- The Committee requested clarification of the 6-point measure scale. Based on input from an expert panel and comparison of current tools in use for similar purposes, the scale proposed was deemed the best fit for purpose. This became important because there is another tool in use by IRFs – the FIM– that is required for payment and uses a different scale; members indicated that facilities may find that confusing if there were different requirements for different programs. CMS indicated that a determination has not been made to convert to function items from the CARE Item Set [tool].

---

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **H-0**; **M-7**; **L-2**; **I-6** (consensus not reached) 2b. Validity: **H-1**; **M-7**; **L-1**; **I-6** (consensus not reached)

UPDATED VOTES FOR 2a. Reliability: H-5; **M-10**; **L-3**; **I-0 2b. Validity: H-4**; **M-12**; **L-2**; **I-0**

<u>Rationale</u>:

- As raised with previous measures, the Committee indicated a strong interest in seeing scientific acceptability data at the facility level. A member notes that Crohnbach alphas provided are at the patient level. The developer indicated they could provide facility level error bars on splines for consideration.
- The Committee asked the developer to consider if it would be more accurate to assess change in function between admission and discharge versus coming up with an expected functional level and seeing if it could be achieved. The assumption is that the comparison to an expected score would be more game-able. The developer indicated they use every bit of data they have available and the true intent of the percent of patients measure is for consumer understandability.
- The Committee acknowledged the wealth of data provided on the reliability and validity of the CARE tool. They continued to struggle with lack of data at the facility level. The developer directed the Committee to supplemental information they provided which may have come in after the Committee reviewed each measure. Supplemental information included the relationship between discharge scores and discharge back to the community and between CARE scores and length of stay.
- The Committee noted that there was some data available, specifically generalized estimation equation data that have splines and error bars, and upon submission that data will be extremely helpful.
- NQF staff clarified that this is not a unique situation and as measures become operationalized, more data becomes available and as this is a standing committee, that data will come back to this committee for further review. There is also the

understanding that with the movement toward pay for performance, Committees want more data and NQF is trying to work those issues into the process.

---

4. Feasibility: H-4; M-8; L-3; I-0

(4a. Clinical data generated during care delivery; 4b. Electronic sources; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)

Rationale:

- The Committee had no questions or concerns on the feasibility of this measure

---

3. Use and Usability: H-3; M-7; L-3; I-2

(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting/Accountability and 3b. Quality Improvement)

Rationale:

- The Committee had no questions or concerns on the use and usability of this measure

---

5. Related and Competing Measures

- The Committee considered this measure to potentially compete with 2286: Functional Change: Change in Self-Care Score (UDSMR) and was asked to vote on these measures. While the Committee agreed that these measures are competing, they did not achieve consensus on whether one measure was superior. When measures 2286 and 2633 moved forward, the CSAC voted to recommend 2286, but not 2633, as the measure was deemed as competing with 2286. The Board of Directors reviewed the recommendations of the CSAC and the rationale for non-approval of 2633. The Board provided greater policy context, including the importance of the IMPACT Act of 2014 and the need for aligned measures that can be used to assess care across settings. The Board therefore directed NQF staff to return the competing IRF measures (2286 and 2633) back to the CSAC for further consideration. In addition, the Board expressed concerns regarding measures derived from proprietary versus non-proprietary instruments, and the desirability of having measures that help assess quality improvement from the patient's perspective as he/she moves among multiple sites of care. In their reconsideration vote, 92% of the CSAC voted to approve endorsement for both measures with conditions for specific update requirements.
- The Committee also considered this measure to be related to 2635: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients (CMS), however, there were no recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-10**; **N-5; UDPATED Y-16**; **N-2**

6. Public and Member Comment: March 2, 2015- March 31, 2015

Comments received:

- One commenter noted that these are important measures but they need to be analyzed and improved as additional data is collected. Another commenter concurred with the Committee's concern with the validity and reliability of measures developed using a cross-sectional study design from a demonstration project, which did not follow the same patients across venues of care and thus limiting applicability across sites.

Committee response:

- The Committee requested additional information to allow for more comprehensive evaluation of the consensus not reached and not recommended measures. This additional information was discussed on the post-comment committee call and the Committee had an opportunity to re-vote on the applicable measures. This measure was recommended by the Committee after reviewing the additional information and the comments.

Developer response:

- Thank you for your comment. As discussed during the measure review on January 22, 2015 and documented in the Person- and Family-Centered Care Phase 2 Draft Report on page 11, the Post-Acute Care Payment Reform demonstration was a prospective cohort study, not a cross-sectional study. In addition to collecting admission and discharge data using the CARE Tool during the post-acute care stay, inpatient claims data for acute care stays prior to and following the post-acute care stay were linked to the CARE admission and discharge data. The reliability and validity of the CARE function items were presented and discussed during the January 21-22, 2015 meeting, and several committee members referred to our analysis as very good. We have also submitted provider-level reliability data to the committee for review, as requested during the January 21-22, 2015 meeting. The Improving Medicare Post Acute Care Transformation (IMPACT) Act directs the Secretary to specify quality measures on which PAC providers are required to submit standardized patient assessment data and other necessary data specified by the Secretary with respect to five quality domains, one of which is functional status, cognitive function, and changes in function and cognitive function.
- The Post-Acute Care Payment Reform Demonstration was a prospective cohort study. It was not a cross-sectional study. For the study, data were collected at admission and discharge for each patient in the study. In addition, we collected interim assessment data for patients in the cost-resource utilization segment of the study. As part of the study, we also linked the CARE admission and discharge data with acute care and post-acute care claims data in order to examine episodes of care and post discharge readmissions. (B). The items and the summed self-care and mobility scores are statistically significantly associated with several outcomes, including length of stay and discharge destination. The admission IRF self-care and IRF mobility scores were moderately correlated with length of stay with coefficients of -0.463 ($p < .0001$) for self-care and -0.474 ($p < .001$) for mobility. As expected, the summed self-care and mobility discharge scores for patients who were discharged to home were significantly different than the scores of patients discharged to a long-term care/nursing home setting. The mean (standard deviation) discharge self-care score for patients going home and to long-term care/nursing home were 34.29 (7.04) and 24.57 (9.39), respectively. For mobility, the mean (standard deviation) scores were 57.35 (15.68)

and 36.57 (15.07), respectively. The patients going home had higher scores, indicating more function, as we expected. (C). The CARE function items included in the 4 IRF quality measures and 2 LTCH quality measures have undergone validity testing. In addition to the results we present in our testing documentation, the data presented above (in 3b), we examined the relationship between the current functional assessment items and the CARE items for each PAC setting. The reports describing the testing are available at: http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/CARE-Item-Set-and-B-CARE.html.

7. Review 1: Consensus Standards Approval Committee (CSAC) Vote June 29, 2015: Y-10; N-5; A-0

Review 2: CSAC Vote September 17, 2015: Y-13; N-0; A-0

CSAC Decision: Approved for endorsement – Final Decision made September 17, 2015

8. Review 1: Board of Directors Review: No (July 22, 2015)

Board Decision: The Board decided to send Measure 2633 back to the CSAC for further consideration.

Review 2: Board of Directors Review: Yes (November 4, 2015)

Board Decision: Ratified for endorsement, with conditions for updates– Final Decision made November 4, 2015

9. Appeals

2634 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients

Submission

Description: This measure estimates the mean risk-adjusted mean change in mobility score between admission and discharge for Inpatient Rehabilitation Facility (IRF) Medicare patients.

Numerator Statement: The measure does not have a simple form for the numerator and denominator. This measure estimates the risk-adjusted change in mobility score between admission and discharge among Inpatient Rehabilitation Facility (IRF) patients age 21 and older. The change in mobility score is calculated as the difference between the discharge mobility score and the admission mobility score.

Denominator Statement: Inpatient Rehabilitation Facility patients included in this measure are at least 21 years of age, Medicare beneficiaries, are not independent with all of the mobility activities at the time of admission, and have complete stays.

Exclusions: This quality measure has 5 exclusion criteria:

1) Patients with incomplete stays.

Rationale: It can be challenging to gather accurate discharge functional status data for patients who experience incomplete stays. Patients with incomplete stays include patients who are unexpectedly discharged to an acute care setting (Short-stay Acute Hospital, Critical Access Hospital, Inpatient Psychiatric Facility, or Long-term Care Hospital) because of a medical emergency; patients who die or leave an Inpatient Rehabilitation Facility (IRF) against medical advice; and patients with a length of stay less than 3 days.

2) Patients who are independent with all mobility activities at the time of admission.

Rationale: Patients who are independent with CARE mobility items at the time of admission are assigned the highest score on all the mobility items, and thus, would not be able to show functional improvement on this same set of items at discharge.

3) Patients with the following medical conditions: coma, persistent vegetative state; complete tetraplegia; locked-in syndrome or severe anoxic brain damage, cerebral edema or compression of brain.

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

4) Patients younger than age 21.

Rationale: There is only limited evidence published about functional outcomes for individuals younger than 21.

5) Patients discharged to hospice.

Rationale: Patient goals may change during the IRF stay.

6) Patients not covered by the Medicare program.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Post Acute/Long Term Care Facility : Inpatient Rehabilitation Facility

Type of Measure: Outcome

Data Source: Electronic Clinical Data

Measure Steward: Centers for Medicare & Medicaid Services (CMS)

STANDING COMMITTEE MEETING [01/21/2015-01/22/2015]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

(1a. Evidence, 1b. Performance Gap, 1c. High Impact)

1a. Evidence: **Y-13**; **N-0**; 1b. Performance Gap: **H-3**; **M-8**; **L-2**; **I-0;** 1c. High Priority: **H-7**; **M-6**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer noted that IRF measures are limited to Medicare only and that the Long-Term Care Hospital Quality Reporting Program was established as a Medicare program. The Committee highlighted that there are talks about these quality measures becoming pay-for-performance measures; however, in IRFs there are currently requirements for pay for performance such as a two-percent reduction in payments for failure to submit certain quality data. The Committee questioned the connection between these specific measures and pay-for-performance measures. The developer clarified that the Inpatient

Rehabilitation Quality Reporting Program assigns a penalty for failure to report, however it is not tied to a pay-for-performance program.

---

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **H-0**; **M-10**; **L-0**; **I-3** 2b. Validity: **H-1**; **M-9**; **L-1**; **I-2**

<u>Rationale</u>:

- The developers utilized different types of reliability including Inter-rater reliability and patient videos reliability. Items that did not test well during the PAC demo were not included. Test-retest reliability was not performed due to the instability of the patients' function.
- The Committee expressed concerns that reliability and validity data was at the care level and not at the facility level; however, since this is an outcome measure the Committee agreed that both reliability and validity should be considered moderate.
- The developers confirmed that the data elements they are using in the risk adjustment model and that the observed or expected calculation comes from the assessment data and comorbidities from the claims data.

---

4. Feasibility: H-6; M-5; L-2; I-0

(4a. Clinical data generated during care delivery; 4b. Electronic sources; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)

<u>Rationale</u>:

- The Committee questioned the length of time it takes to administer or grade the instrument. The developer noted that clinicians are assessing patients on the ability to complete the activities listed in the measure.

---

3. Use and Usability: H-6; M-5; L-0; I-2

(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting/Accountability and 3b. Quality Improvement)

<u>Rationale</u>:

- The Committee had no concerns with the usability of the measure.

---

5. Related and Competing Measures

- The Committee considered this measure to compete with 2321: Functional Change: Change in Mobility Score (UDSMR). While the Committee agreed that these measures are competing, they did not achieve consensus on whether one measure was superior. When measures 2321 and 2634 moved forward, the CSAC voted to recommend 2321, but not 2634, as the measure was deemed as competing with 2286. The Board of Directors reviewed the recommendations of the CSAC and the rationale for non-approval of 2634. The Board provided greater policy context, including the importance

of the IMPACT Act of 2014 and the need for aligned measures that can be used to assess care across settings. The Board therefore directed NQF staff to return the competing IRF measures (2321 and 2634) back to the CSAC for further consideration. In addition, the Board expressed concerns regarding measures derived from proprietary versus non-proprietary instruments, and the desirability of having measures that help assess quality improvement from the patient's perspective as he/she moves among multiple sites of care. In their reconsideration vote, 92% of the CSAC voted to approve endorsement for both measures with conditions for specific update requirements.

- The Committee also considered this measure to be related to 2636: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Mobility Score for Medical Rehabilitation Patients (CMS), however there were no recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-11**; **N-2**

6. Public and Member Comment: March 2, 2015- March 31, 2015

Comments received:

- Measures 2634 and 2636 received two similar comments. The first commenter supported the underlying concept of the measures, stating that inpatient rehabilitation facilities need to be measured on outcomes based on functional improvement. However, the commenter suggested that an alternative measure that determines how the provider improved the patient's life (mobility) would better incentivize a change in clinical practice and associated patient-level outcomes as opposed to measure 2634 and measure 2636. Another commenter concurred with the Committee's concern with the validity and reliability of measures developed using a cross-sectional study design from a demonstration project, which did not follow the same patients across venues of care and thus limiting applicability across sites.

NQF response:

- NQF is limited to reviewing measures that are submitted for endorsement. We have added this suggestion to the measure gap list in the report. Thank you for your comment.

Developer response:

- Thank you for your comment. As discussed during the measure review on January 22, 2015 and documented in the Person- and Family-Centered Care Phase 2 Draft Report on page 11, the Post-Acute Care Payment Reform demonstration was a prospective cohort study, not a cross-sectional study. In addition to collecting admission and discharge data using the CARE Tool during the post-acute care stay, inpatient claims data for acute care stays prior to and following the post-acute care stay were linked to the CARE admission and discharge data. The reliability and validity of the CARE function items were presented and discussed during the January 21-22, 2015 meeting, and several committee members referred to our analysis as very good. We have also submitted provider-level reliability data to the committee for review, as requested during the January 21-22, 2015 meeting. The Improving Medicare Post Acute Care Transformation (IMPACT) Act directs the Secretary to specify quality measures on which PAC providers

are required to submit standardized patient assessment data and other necessary data specified by the Secretary with respect to five quality domains, one of which is functional status, cognitive function, and changes in function and cognitive function.

---

7. Review 1: Consensus Standards Approval Committee (CSAC) Vote June 29, 2015: Y-10; N-5; A-0

Review 2: CSAC Vote September 17, 2015: Y-13; N-0; A-0

CSAC Decision: Approved for endorsement – Final Decision made September 17, 2015

8. Review 1: Board of Directors Review: No (July 22, 2015)

Board Decision: The Board decided to send Measure 2634 back to the CSAC for further consideration.

Review 2: Board of Directors Review: Yes (November 4, 2015)

Board Decision: Ratified for endorsement, with conditions for updates– Final Decision made November 4, 2015

---

9. Appeals

Appendix E: Board Memo

TO: The NQF Board of Directors
FR: Helen Burstin, Chief Scientific Officer
Marcia Wilson, Senior Vice President, Quality Measurement
RE: Ratification of Measures for the Person- and Family-Centered Care Phase 2 Project

DA: October 28, 2015

ACTION REQUESTED
The Board of Directors is asked to ratify the CSAC's recommendation to endorse four measures for the second phase of the Person- and Family-Centered Care (PFCC) project. It is recommended that the measures be endorsed with special update requirements for the following four measures from both measure stewards. (See Appendices A, B and C for additional measure level detail.)

- 2633: IRF Functional Outcome Measure: Change in Self-Care Score for Medical Rehabilitation Patients (CMS)
- 2634: IRF Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients (CMS)
- 2286: Functional Change: Change in Self Care Score (UDSMR)
- 2321: Functional Change: Change in Mobility Score (UDSMR)

BACKGROUND

Upon request for re-consideration of the above four measures, the CSAC recommends approval with the conditions stated below. The Board of Directors reviewed the recommendations of the CSAC and the rationale for non-approval of two of the measures. The Board provided greater policy context including the importance of the IMPACT Act enacted in 2014 and the need for aligned measures that can be used to assess care across settings. The Board therefore directed NQF staff to return the four competing IRF measures in question back to the CSAC for further consideration. (See Appendix B for a side by side comparison of the competing measures.). In addition, the Board expressed concerns regarding measures derived from proprietary versus non-proprietary instruments, and the desirability of having measures that help assess quality improvement from the patient's perspective as he/she moves among multiple sites of care

As was true at the June CSAC meeting, there were extensive public comments made during the Board meeting. The FIM tool proponents primarily focused on concerns around the sensitivity of the CARE tool measures, the burden of having to report on two sets of measures for the same setting and the concerns about having to use a new tool (CARE) after providers have built considerable infrastructure (e.g., staff training, software) to collect data with the FIM Instrument. The CARE tool proponents supported measures developed from the CARE tool because they recognize the importance of all providers moving to just one tool and they supported the CMS' decision to use the CARE tool across multiple settings. Additionally, proponents supported the use of a non-proprietary assessment tool generally.

Consensus Process to Date:

These four measures were recommended for endorsement by the Standing Committee after considerable public comment, member voting and additional information provided by measure developers. (Appendix C provides themes from the public comments.) The Standing Committee was unable to select a best-in-class for either set of competing measures (#2633 versus 2286 and #2634 versus 2321). The two UDSMR measures (#2286 and 2321) were recommended for endorsement with 71% of councils approving. The councils were unable to reach consensus for the two CMS measures (#2633 and 2634) with only 56% of councils approving the measures.

In their initial vote in June, the CSAC voted to recommend the two UDSMR measures, while the two CARE tool measures only received 56% approval (below the required 60% threshold for CSAC approval). Based on the rationale provided by CSAC members, the CMS IRF measures were not approved largely due to competing measure concerns.

In their reconsideration vote in September, **12 out of 13 CSAC members or 92% voted to Approve endorsement for the four measures with conditions for specific updates**

Update Requirements:

UDSMR	**CMS**
• Provide information about how the inclusion or exclusion of cognitive items impacts the overall assessment of the patient. • Provide updated measure level testing for reliability and validity given that all the measures are new. There is particular interest in measure performance/scientific acceptability across care settings beyond IRF. • Provide information about costs associated with use of the FIM Instrument, respective software and tools; and costs of ongoing training in order to accurately use the FIM Instrument.	• Provide information about how the inclusion or exclusion of cognitive items impacts the overall assessment of the patient. • Provide updated measure level testing for reliability and validity given that all the measures are new and will be implemented in 2016. • Provide data on comparison of the competing measure results to gain an understanding of which scale is more reliable, valid and feasible. • Provide a summary of qualitative data gathered during rule-making process including perceived benefits from the field for instruments that cut across settings.

Appendix A: Additional Measure Level Detail for Four Candidate Consensus Standards

Measure	Steward	Committee Recommendation and Member Votes for Approval	Type of Measure	Measure* Setting of Care - Level of Analysis	Assessment tool Used	Standing Committee History/Considerations
2286: Functional Change: Change in Self Care Score (new)	UDSMR	Committee: Recommended % Councils Approving: 71% % CSAC Approving (original vote): 100%	Outcome	IRF – Facility	FIM® Instrument	Measure recommended at In-Person Meeting; while additional information not required, the Committee requested disparities data (data for race, age, payer); intra-class co-efficient at the facility level; and mean fit statistics.
2321: Functional Change: Change in Mobility Score (Uniform Data System for Medical Rehabilitation) (new)	UDSMR	Committee: Recommended % Councils Approving: 71% % CSAC Approving (original vote): 100%	Outcome	IRF – Facility	FIM® Instrument	Measure recommended at In-Person Meeting; while additional information not required, the Committee requested disparities data (data for race, age, payer); intra-class co-efficient at the facility level; and mean fit statistics.
2633: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Self- Care Score for Medical Rehabilitation Patients (new)	CMS	Committee: Recommended % Councils Approving: 57% % CSAC Approving (original vote): 56%	Outcome	IRF – Facility	CARE Item Set	Consensus Not Reached on Reliability and Validity at In-Person Meeting. Additional information was provided on reliability, validity and performance at the facility level and the Committee subsequently recommended the measure for endorsement.
2634: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation	CMS	Committee: Recommended % Councils Approving: 57% % CSAC Approving (original vote): 56%	Outcome	IRF– Facility	CARE Item Set	Measure recommended at In-Person Meeting. No additional information requested from the developer for clarification of NQF criteria.

*Note: While the assessment tools (or item sets) used to calculate these measures may be used in more than one setting, the Standing Committee evaluated and recommended endorsement based on the MEASURE submission form and information provided in the measure description, evidence, rationale, etc. As with the measures submitted for specific settings utilizing the CARE Item Set, UDSMR has been advised to prepare new measure submissions for settings of care beyond IRFs for the FIM tool.

Appendix C: Themes from Public Comment

In addition to the two sets of competing measures, the CSAC also voted on eight additional measures. Out of these twelve measures, three were derived from the FIM® Instrument for use in an Inpatient Rehabilitation Facility and nine were derived from the CARE tool for use in different settings including Inpatient Rehabilitation Facility. There were a number of comments received during the Public Comment period on the June 9th CSAC call. Many of the comments covered issues that had previously been raised either by the Standing Committee during measure evaluation or during the Public Comment period, and can be summarized as follows:

1. **Sensitivity of the CARE tool:** The overarching concern from the provider community is that the CARE tool is not sensitive enough to assess improvement in patients, and with this lack of sensitivity at the patient level there was question about impact on the overall measure. The Standing Committee conducted a detailed review of data at both the scale/item level and subsequently at various facility levels for each of the measures, regardless of the assessment tool used. They did not perceive a concern with the sensitivity testing conducted at the CARE item set. CMS and their measure development contractors re-iterated substantial testing at both levels of analysis (item and facility) that indicated the ability to discriminate between facilities. Additional measures based on the CARE Item Set, but developed by the American Health Care Association (AHCA), were supported by data that demonstrated sensitivity at both the item and facility levels. The measure developers have provided detailed responses on this issue in the attached memos. CMS response is located at this link and AHCA response can be found at the following link.
2. **Measurement Burden**: As indicated above, having multiple measures with the same focus and designed for the same care settings is expected to cause substantial burden on facility staff; this was a consideration by the Standing Committee and is part of the rationale for inability to reach consensus regarding harmonization or determination of best in class measures. The discussion around burden of measurement centered around the collection of the following assessment tools/item sets:
    a. The FIM System® is a an outcomes management program for skilled nursing facilities, sub-acute facilities, long-term care hospitals, Veterans Administration programs, international rehabilitation hospitals, and other related venues of care. While the FIM® has been collected for some time, the measures submitted for this project (#2286, 2287 and 2321) are considered new for endorsement. It should also be noted that the measures submitted, while potentially applicable for additional settings, were only considered for IRFs. The measure submission forms, including measure titles, descriptions, rationale and evidence provided were specific to IRFs, thus the Committee was directed to only consider that setting.
    b. CARE Item Set: As a part of the Medicare Post-Acute Care Payment Reform Demonstration (PAC-PRD), a standardized patient assessment tool was developed for use at acute hospital discharge and at post-acute care admission and discharge. This tool was named the Continuity Assessment Record and Evaluation (CARE) Item Set. Data collected using the CARE Item Set served as a major source of information in the demonstration. The CARE Item Set measures the health and functional status of Medicare beneficiaries at acute discharge, and measures changes in severity and other outcomes for Medicare post-acute care patients. The CARE Item Set is designed

to standardize assessment of patients' medical, functional, cognitive, and social support status across acute and post-acute settings, including long-term care hospitals (LTCHs), inpatient rehabilitation facilities (IRFs), skilled nursing facilities (SNFs), and home health agencies (HHAs). The goal was to standardize the items used in each of the existing assessment tools while posing minimal administrative burden to providers. Nine (9) measures were submitted to this project, and are based on data derived from use of the CARE tool.

3. **Measure Gaps**: The Standing Committee and public comments expressed the need and interest in measures that focus on patient stabilization, when improvement is not the goal of treatment; and also for measures more directly related to patient goals versus treatment goals.