Appendix A: Person- and Family-Centered Care Phase 2 Measure Evaluation Summary Tables

LEGEND: Y = Yes; N = No; H = High; M = Moderate; L = Low; I = Insufficient

0167 Improvement in Ambulation/locomotion

Submission |

Description: Percentage of home health episodes of care during which the patient improved in ability to ambulate.

Numerator Statement: Number of home health episodes of care where the value recorded on the discharge assessment indicates less impairment in ambulation/locomotion at discharge than at start (or resumption) of care.

Denominator Statement: Number of home health episodes of care ending with a discharge during the reporting period, other than those covered by generic or measure-specific exclusions.

Exclusions: All home health episodes where the value recorded for the OASIS-C item M1860 ("Ambulation/Locomotion") on the start (or resumption) of care assessment indicates minimal or no impairment, or the patient is non-responsive, or the episode of care ended in transfer to inpatient facility or death at home, or the episode is covered by the generic exclusions.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Home Health

Type of Measure: Outcome

Data Source: Electronic Clinical Data

Measure Steward: Centers for Medicare & Medicaid Services

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-14; N-3**; 1b. Performance Gap: **H-3**; **M-13**; **L-1**; **I-0;** 1c. High Priority: **H-10**; **M-7**; **L-0**; **I-0**

Rationale:

- Overall the Committee felt that this is an important indicator to assess improvement in a patient's functional status in performing activities of daily living, which would allow patients to remain in their home environment rather than moving to a facility-based setting. However, the Committee had a major concern that this measure set focuses on "improvement" in mobility when the Jimmo v. Sebelius settlement requires CMS to not require improvement in function as a condition of coverage in home health (as well as SNF and outpatient services). The Committee expressed concern that by endorsing an ambulation measure that evaluates improvement, home health agencies may be more likely to deny access to patients who require home health services to maintain or prevent further deterioration of function, but have no realistic potential to improve. The Committee recommended these patients should be added to the denominator exclusions so not to create a system where there are disincentives to treat people who may not improve but still might need that therapy in order to maintain or prevent deterioration of function.
- CMS noted they agree with the Committee's concern and indicated that they are moving in that direction where they are able to balance out the incentives so that they are not developing measures that are seen as potentially disincentivizing care or taking away from the actual goals of the resident or the patient.

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

0167 Improvement in Ambulation/locomotion

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)
2a. Reliability: **H-6**; **M-8**; **L-1**; **I-2** 2b. Validity: **H-5**; **M-7**; **L-2**; **I-3**
Rationale:

- Testing was conducted at both the patient level and the organizational level. Beta Binomial Reliability testing was conducted and considered above the range for reliability. Also, acceptable test-retest reliability was shown by the data, suggesting the test is repeatable and yields consistent results. Measure through OASIS achieved substantial inter-rater reliability.
- Both patient and organizational levels were tested and validated through a variety of approaches. Validity testing was comprehensive and included consensus validity by experts, convergent predictive validity, and validation by outcome enhancement. Data demonstrated statistically significant correlations between the measure and improvement in outcomes/quality. Data also demonstrated widespread implementation of this measure is appropriate

4. Feasibility: H-9; M-7; L-1; 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:

- OASIS data collection and transmission is a requirement for the Medicare Home Health Conditions of Participation. Information on ambulation status used to calculate this measure is recorded in the relevant OASIS items embedded in the agency´s clinical assessment as part of normal clinical practice. OASIS data are collected by the home health agency during the care episode and transmitted electronically to the state and CMS national OASIS repository.

3. Use and Usability: H-4; M-11; L-2; I-0
(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting/Accountability and 3b. Quality Improvement)
Rationale:

- This measure is already widely used and publicly reported in a variety of places, including Home Health Compare and CMS´s Home Health Quality Initiative "Outcome Quality Measure Report", which provides all Medicare-certified home health agencies with opportunities to use outcome measures for outcome-based quality improvement. The report allows agencies to benchmark their performance against other agencies across the state and nationally, as well as their own performance from prior time periods.

5. Related and Competing Measures

- The Committee considered measures 0167, 0174, and 0175 to be related to measure 2287 (Functional Change: Change in Motor Score), as they have the same focus area but were considered for different settings and populations. Measures 0167, 0174, and 0175 are intended for certified home health patients 18 and above and measure 2287 was submitted as an inpatient rehabilitation facility measure. The Committee agreed that there was a need for all of the aforementioned measures; they made no recommendations for harmonization. In alignment with the IMPACT Act provisions, the Committee emphasized the importance of using cross-setting measures in programs for future considerations.

Standing Committee Recommendation for Endorsement: **Y-17**; **N-0**

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

- Measures 0167, 0174, and 0175 received two sets of comments suggesting that they be combined to be a composite that would “collectively address daily living activities”. In addition, it was suggested that the

0167 Improvement in Ambulation/locomotion

measure specifications be revised to “measure patients upon meeting expected outcomes of interventions versus the achievement of patient goals”.

- Three comments were received regarding the exclusions, one in favor, and two raising concerns: that these measures may discriminate against patients who require therapy to maintain abilities, but who may not improve due to their condition, and that not enough rare/unpredictable diseases could be included, therefore leading to potential unintended consequences.

Committee response:

- The Committee had discussed this issue during the in-person meeting, especially raising concerns around the Jimmo v. Sebelius settlement, but felt that the developer had adequately addressed it using the exclusion criteria. Additionally, the developer (CMS) noted that they are working to balance incentives to ensure patients who will not improve are still receiving high quality care and are not being discriminated against.

Developers response:

- Thank you for your comment. The current OASIS does not allow for the inclusion of patient goals in the calculation of a measure. In the context of implementing cross-setting measures, some items related to goals for patient functioning may be added to the assessment and could be the basis for additional quality measures. CMS is also exploring composite functional measures for future development.
- Thank you for your comment. We recognize that there are some home health patients for whom improvement in ambulation/locomotion is not a reasonable expectation. Risk adjustment, while not perfect, helps to mitigate the effect of the patient's clinical condition at admission and other patient characteristics on the home health agency's measure value. Notwithstanding recent changes in the types of patients accepted for home health care, it remains primarily a post-acute benefit. The measure steward will continue to explore options for refining the measure based on committee input and comments received, and will explore potential alternative measures that address ambulation/locomotion outcomes for patients with limited likelihood of improvement.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0174 Improvement in bathing

Submission |

Description: Percentage of home health episodes of care during which the patient got better at bathing self.

Numerator Statement: Number of home health episodes of care where the value recorded on the discharge assessment indicates less impairment in bathing at discharge than at start (or resumption) of care.

Denominator Statement: Number of home health episodes of care ending with a discharge during the reporting period, other than those covered by generic or measure-specific exclusions.

Exclusions: All home health episodes where at the start (or resumption) of care assessment the patient had minimal or no impairment, or the patient is non-responsive, or the episode of care ended in transfer to inpatient facility or death at home, or was covered by the generic exclusions.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Home Health

Type of Measure: Outcome

Data Source: Electronic Clinical Data

0174 Improvement in bathing

Measure Steward: Centers for Medicare & Medicaid Services

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-14; N-3**; 1b. Performance Gap: **H-3**; **M-13**; **L-1**; **I-0;** 1c. High Priority: **H-10**; **M-7**; **L-0**; **I-0**

<u>Rationale</u>:

- The Committee questioned the gap between the measured outcome and the evidence to support those interventions that would support improvement, noting there are no practice guidelines around educating people on bathing. However, the Committee agreed that this is an important indicator because the goal is that home health patients to be independent and able to have the ability to bathe themselves.
- CMS indicated that this type of outcome measure was created so that CMS can benchmark, set thresholds, or publically report in that setting.
- The Committee's remarks on measure 0167 regarding the Jimmo v. Sebelius settlement, which requires CMS to not require improvement in function as a condition of coverage in home health (as well as SNF and outpatient services) would apply to all measures addressing improvement in ADLs.

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-8**; **L-1**; **I-2** 2b. Validity: **H-5**; **M-7**; **L-2**; **I-3**

<u>Rationale</u>:

- The Committee commented that all reliability testing indicates measure reliability and inter-rater reliability was high for this measure and a large number of cases were sampled.
- The Committee also noted that the validity testing is also high for this measure and consistent with evidence based practice.

4. Feasibility: H-9; M-7; L-1; 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 commented that every patient that meets the eligibility criteria for performing the initial OASIS assessment is included and that the required data elements are all routinely generated and used during care delivery. The data collection strategy is already operationalized.

3. Use and Usability: H-4; M-11; L-2; I-0

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

<u>Rationale</u>:

- The measure is already widely used and publicly reported in a variety of places, including Home Health Compare and CMS´s Home Health Quality Initiative "Outcome Quality Measure Report", which provides all Medicare-certified home health agencies with opportunities to use outcome measures for outcome-based quality improvement. The report allows agencies to benchmark their performance against other agencies across the state and nationally, as well as their own performance from prior time periods.

5. Related and Competing Measures

- The Committee considered measures 0167, 0174, and 0175 to be related to measure 2287 (Functional

0174 Improvement in bathing

Change: Change in Motor Score), as they have the same focus area but considered for different settings and populations. Measures 0167, 0174, and 0175 are intended for certified home health patients 18 and above and measure 2287 was submitted as an inpatient rehabilitation facility measure. The Committee agreed that there was a need for all of the aforementioned measures; they made no recommendations for harmonization. In alignment with the IMPACT Act provisions, the Committee emphasized the importance of using cross-setting measures in programs for future considerations.

Standing Committee Recommendation for Endorsement: **Y-17**; **N-0**

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

Comments received:

- Measures 0167, 0174, and 0175 received two sets of comments suggesting that they be combined to be a composite that would "collectively address daily living activities". In addition, it was suggested that the measure specifications be revised to "measure patients upon meeting expected outcomes of interventions versus the achievement of patient goals".
- Three comments were received regarding the exclusions, one in favor, and two raising concerns: that these measures may discriminate against patients who require therapy to maintain abilities, but who may not improve due to their condition, and that not enough rare/unpredictable diseases could be included, therefore leading to potential unintended consequences.

Committee response:

- The Committee had discussed this issue during the in-person meeting, especially raising concerns around the Jimmo v. Sebelius settlement, but felt that the developer had adequately addressed it using the exclusion criteria. Additionally, the developer (CMS) noted that they are working to balance incentives to ensure patients who will not improve are still receiving high quality care and are not being discriminated against.

Developers response:

- Thank you for your comment. The current OASIS does not allow for the inclusion of patient goals in the calculation of a measure. In the context of implementing cross-setting measures, some items related to goals for patient functioning may be added to the assessment and could be the basis for additional quality measures. CMS is also exploring composite functional measures for future development.
  Thank you for your comment. We recognize that there are some home health patients for whom improvement in ambulation/locomotion is not a reasonable expectation. Risk adjustment, while not perfect, helps to mitigate the effect of the patient's clinical condition at admission and other patient characteristics on the home health agency's measure value. Notwithstanding recent changes in the types of patients accepted for home health care, it remains primarily a post-acute benefit. The measure steward will continue to explore options for refining the measure based on committee input and comments received, and will explore potential alternative measures that address ambulation/locomotion outcomes for patients with limited likelihood of improvement.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0175 Improvement in bed transferring

Submission |

Description: Percentage of home health episodes of care during which the patient improved in ability to get in and out of bed.

0175 Improvement in bed transferring

Numerator Statement: Number of home health episodes of care where the value recorded on the discharge assessment indicates less impairment in bed transferring at discharge than at start (or resumption) of care.

Denominator Statement: Number of home health episodes of care ending with a discharge during the reporting period, other than those covered by generic or measure-specific exclusions.

Exclusions: All home health episodes where at the start (or resumption) of care assessment the patient is able to transfer independently, or the patient is non-responsive. or the episode of care ended in transfer to inpatient facility or death at home, or the episode is covered by the generic exclusions.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Home Health

Type of Measure: Outcome

Data Source: Electronic Clinical Data

Measure Steward: Centers for Medicare & Medicaid Services

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-14; N-3**; 1b. Performance Gap: **H-3**; **M-13**; **L-1**; **I-0;** 1c. High Priority: **H-10**; **M-7**; **L-0**; **I-0**

<u>Rationale</u>:

- The Committee commented that the measure directly applies to the function that is being measured and there is a demonstrated and documented performance gap.
- The Committee's remarks on measure 0167 regarding the Jimmo v. Sebelius settlement which requires CMS to not require improvement in function as a condition of coverage in home health (as well as SNF and outpatient services) apply to all measures addressing improvement in ADLs.

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-8**; **L-1**; **I-2** 2b. Validity: **H-5**; **M-7**; **L-2**; **I-3**

<u>Rationale</u>:

- The Committee commented that testing is consistent with target population of Medicare consumers, but not representative of other consumers needing Home Health Services.

4. Feasibility: H-9; M-7; L-1; 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>:

- OASIS data collection and transmission is a requirement for the Medicare Home Health Conditions of Participation. Information on bed transferring status used to calculate this measure is recorded in the relevant OASIS items embedded in the agency´s clinical assessment as part of normal clinical practice. OASIS data are collected by the home health agency during the care episode and transmitted electronically to the state and CMS national OASIS repository.

3. Use and Usability: H-4; M-11; L-2; I-0

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

<u>Rationale</u>:

- The measure is already widely used and publicly reported in a variety of places, including Home Health

0175 Improvement in bed transferring
Compare and CMS´s Home Health Quality Initiative "Outcome Quality Measure Report", which provides all Medicare-certified home health agencies with opportunities to use outcome measures for outcome-based quality improvement. The report allows agencies to benchmark their performance against other agencies across the state and nationally, as well as their own performance from prior time periods.
5. Related and Competing Measures • The Committee considered measures 0167, 0174, and 0175 to be related to 2287 (Functional Change: Change in Motor Score), as they have the same focus area but were considered for different settings and populations. Measures 0167, 0174, and 0175 are intended for certified home health patients 18 and above and measure 2287 was submitted as an inpatient rehabilitation facility measure. The Committee agreed that there was a need for all of the aforementioned measures; they made no recommendations for harmonization. In alignment with the IMPACT Act provision, the Committee emphasized the importance of using cross-setting measures in programs for future considerations.
Standing Committee Recommendation for Endorsement: **Y-17**; **N-0**
6. Public and Member Comment: March 2, 2015- March 31, 2015 Comments received: • Measures 0167, 0174, and 0175 received two sets of comments suggesting that they be combined to be a composite that would “collectively address daily living activities”. In addition, it was suggested that the measure specifications be revised to “measure patients upon meeting expected outcomes of interventions versus the achievement of patient goals”. • Three comments were received regarding the exclusions, one in favor, and two raising concerns: that these measures may discriminate against patients who require therapy to maintain abilities, but who may not improve due to their condition, and that not enough rare/unpredictable diseases could be included, therefore leading to potential unintended consequences. Committee response: • The Committee had discussed this issue during the in-person meeting, especially raising concerns around the Jimmo v. Sebelius settlement, but felt that the developer had adequately addressed it using the exclusion criteria. Additionally, the developer (CMS) noted that they are working to balance incentives to ensure patients who will not improve are still receiving high quality care and are not being discriminated against. Developers response: • Thank you for your comment. The current OASIS does not allow for the inclusion of patient goals in the calculation of a measure. In the context of implementing cross-setting measures, some items related to goals for patient functioning may be added to the assessment and could be the basis for additional quality measures. CMS is also exploring composite functional measures for future development. • Thank you for your comment. We recognize that there are some home health patients for whom improvement in ambulation/locomotion is not a reasonable expectation. Risk adjustment, while not perfect, helps to mitigate the effect of the patient's clinical condition at admission and other patient characteristics on the home health agency's measure value. Notwithstanding recent changes in the types of patients accepted for home health care, it remains primarily a post-acute benefit. The measure steward will continue to explore options for refining the measure based on committee input and comments received, and will explore potential alternative measures that address ambulation/locomotion outcomes for patients with limited likelihood of improvement.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals

0176 Improvement in management of oral medications

Submission |

Description: Percentage of home health episodes of care during which the patient improved in ability to take their medicines correctly, by mouth.

Numerator Statement: Number of home health episodes of care where the value recorded on the discharge assessment indicates less impairment in taking oral medications at discharge than at start (or resumption) of care.

Denominator Statement: Number of home health episodes of care ending with a discharge during the reporting period, other than those covered by generic or measure-specific exclusions.

Exclusions: All home health episodes where at start (or resumption) of care the patient is not taking any oral medications or has minimal or no impairment, or the patient is non-responsive, or the episode of care ended in transfer to inpatient facility or death, or the episode is covered by the generic exclusions.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Home Health

Type of Measure: Outcome

Data Source: Electronic Clinical Data

Measure Steward: Centers for Medicare & Medicaid Services

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-15**; **N-1**; 1b. Performance Gap: **H-5**; **M-10**; **L-1**; **I-0;** 1c. High Priority: **H-12**; **M-4**; **L-0**; **I-0**

Rationale:

- The Committee commented that this measure directly applies to the process being measured, population disparity has been demonstrated, and there is room for improvement.
- The Committee's remarks on measure 0167 regarding the Jimmo v. Sebelius settlement, which requires CMS to not require improvement in function as a condition of coverage in home health (as well as SNF and outpatient services), applies to all measures addressing improvement in ADLs.

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-2**; **M-9**; **L-0**; **I-5** 2b. Validity: **H-1**; **M-9**; **L-1**; **I-5**

Rationale:

- The Committee commented that the measure is already used in wide-spread implementation by CMS as part of a set of measures for Home Health reporting and testing done at both levels with adequate scope and method.
- The Committee also noted that almost 50% of cases excluded from the measure, but may reflect the general fragility of the population. So the exclusions seem reasonable.

4. Feasibility: H-6; M-9; L-1; 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:

- OASIS data collection and transmission is a requirement for the Medicare Home Health Conditions of Participation. Information on oral medication management status used to calculate this measure is

0176 Improvement in management of oral medications
recorded in the relevant OASIS items embedded in the agency´s clinical assessment as part of normal clinical practice. OASIS data are collected by the home health agency during the care episode and transmitted electronically to the state and CMS national OASIS repository.
3. Use and Usability: H-3; M-9; L-3; I-1 *(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting/Accountability and 3b. Quality Improvement)* Rationale: • The measure is already widely used and publicly reported in a variety of places, including Home Health Compare and CMS´s Home Health Quality Initiative "Outcome Quality Measure Report", which provides all Medicare-certified home health agencies with opportunities to use outcome measures for outcome-based quality improvement. The report allows agencies to benchmark their performance against other agencies across the state and nationally, as well as their own performance from prior time periods.
5. Related and Competing Measures • No related or competing measures noted.
Standing Committee Recommendation for Endorsement: **Y-16**; **N-0**
6. Public and Member Comment: March 2, 2015- March 31, 2015 Comments received: • These two measures received two comments indicating a lack of support. Commenters stated that the methodology used to show improvement is subjective and that the measure does not add value to the portfolio. • Another commenter raised two concerns with this pair of measures, first the potential disincentives for maintenance therapy, and second, the related concern that the list of exclusions is not broad enough. Committee response: • During the in-person meeting, the Committee specifically requested more information on the usability of the two additional concepts (ability to take medicines correctly and frequency of pain) and noted these might be better operationalized via patient reported outcomes due to their subjectivity. However, the Committee voted to recommend these measures for endorsement. • The issues of unintended consequences, “cherry-picking” patients for inclusion in measures, and assessing “improvement” for payment or penalty use in quality programs were discussed during the in-person meeting. The Committee continues to encourage measure developers and implementers to consider implications of measurement, including potential unintended consequences. Developer response: • Thank you for your comment. We recognize that there are some home health patients for whom improvement in management of oral medications is not a reasonable expectation. Risk adjustment, while not perfect, helps to mitigate the effect of the patient's clinical condition at admission and other patient characteristics on the home health agency's measure value. Notwithstanding recent changes in the types of patients accepted for home health care, it remains primarily a post-acute benefit. The measure steward will continue to explore options for refining the measure based on committee input and comments received, and will explore potential alternative measures that address management of oral medications outcomes for patients with limited likelihood of improvement. • Thank you for your comment. Centers for Medicare & Medicaid Services (CMS) will review your comment and address your concerns shortly.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals

0177 Improvement in pain interfering with activity

Submission |

Description: Percentage of home health episodes of care during which the frequency of the patient's pain when moving around improved.

Numerator Statement: Number of home health episodes of care where the value recorded on the discharge assessment indicates less frequent pain at discharge than at start (or resumption) of care.

Denominator Statement: Number of home health episodes of care ending with a discharge during the reporting period, other than those covered by generic or measure-specific exclusions.

Exclusions: All home health episodes where there is no pain reported at the start (or resumption) of care assessment, or the patient is non-responsive, or the episode of care ended in transfer to inpatient facility or death at home, or the episodes is covered by one of the generic exclusions.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Home Health

Type of Measure: Outcome

Data Source: Electronic Clinical Data

Measure Steward: Centers for Medicare & Medicaid Services

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-16**; **N-0**; 1b. Performance Gap: **H-5**; **M-10**; **L-0**; **I-1;** 1c. High Priority: **H-11**; **M-5**; **L-0**; **I-0**

Rationale:

- The Committee noted that pain management is a significant health issue related to functional outcomes and there is definitely a relationship between the measured outcome and a healthcare action supported by the rationale. Pain is extremely important in patient outcomes and needs to be measured accurately.
- The Committee's remarks on measure 0167 regarding the Jimmo v. Sebelius settlement, which requires CMS to not require improvement in function as a condition of coverage in home health (as well as SNF and outpatient services), applies to all measures addressing improvement in ADLs.

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-3**; **M-9**; **L-0**; **I-4** 2b. Validity: **H-2**; **M-8**; **L-2**; **I-4**

Rationale:

- The Committee requested clarification as to whether this measure evaluates frequency of pain, levels of pain, or both. The developer stated that they are not using any scale but evaluating how often pain interferes with activities.
- The Committee stated that there is testing information regarding both data element and scores, and the results demonstrate sufficient reliability.

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

- OASIS data collection and transmission is a requirement for the Medicare Home Health Conditions of Participation. Information on pain interfering with activity used to calculate this measure is recorded in the relevant OASIS items embedded in the agency´s clinical assessment as part of normal clinical practice.

0177 Improvement in pain interfering with activity

OASIS data are collected by the home health agency during the care episode and transmitted electronically to the state and CMS national OASIS repository.

3. Use and Usability: H-5; M-11; L-0; I-0

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

Rationale:

- The measure is already widely used and publicly reported in a variety of places, including Home Health Compare and CMS´s Home Health Quality Initiative "Outcome Quality Measure Report", which provides all Medicare-certified home health agencies with opportunities to use outcome measures for outcome-based quality improvement. The report allows agencies to benchmark their performance against other agencies across the state and nationally, as well as their own performance from prior time periods.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: Y-15; N-1

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

Comments received:

- These two measures received two comments indicating a lack of support. Commenters stated that the methodology used to show improvement is subjective and that the measure does not add value to the portfolio.
- Another commenter raised two concerns with this pair of measures, first the potential disincentives for maintenance therapy, and second, the related concern that the list of exclusions is not broad enough.

Committee response:

- During the in-person meeting, the Committee specifically requested more information on the usability of the two additional concepts (ability to take medicines correctly and frequency of pain) and noted these might be better operationalized via patient reported outcomes due to their subjectivity. However, the Committee voted to recommend these measures for endorsement.
- The issues of unintended consequences, “cherry-picking” patients for inclusion in measures, and assessing “improvement” for payment or penalty use in quality programs were discussed during the in-person meeting. The Committee continues to encourage measure developers and implementers to consider implications of measurement, including potential unintended consequences.

Developer response:

- Thank you for your comment. We recognize that there are some home health patients for whom improvement in management of oral medications is not a reasonable expectation. Risk adjustment, while not perfect, helps to mitigate the effect of the patient's clinical condition at admission and other patient characteristics on the home health agency's measure value. Notwithstanding recent changes in the types of patients accepted for home health care, it remains primarily a post-acute benefit. The measure steward will continue to explore options for refining the measure based on committee input and comments received, and will explore potential alternative measures that address management of oral medications outcomes for patients with limited likelihood of improvement.
- Thank you for your comment. Centers for Medicare & Medicaid Services (CMS) will review your comment and address your concerns shortly.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0422 Functional status change for patients with Knee impairments

Submission | Specifications

Description: A self-report measure of change in functional status for patients 18year+ with knee impairments. The change in functional status assessed using FOTO's (knee) PROM is adjusted to patient characteristics known to be associated with functional status outcomes (risk adjusted) and used as a performance measure at the patient level, at the individual clinician, and at the clinic level to assess quality.

Numerator Statement: Patient Level: The residual functional status score for the individual patient (residual scores are the actual change scores - predicted change after risk adjustment.

Individual Clinician Level: The average of residuals in functional status scores in patients who were treated by a clinician in a 12 month time period for knee impairment.

Clinic Level: The average of residuals in functional status scores in patients who were treated by a clinic in a 12 month time period for knee impairment.

Denominator Statement: All patients 18 years and older with knee impairments who have initiated rehabilitation treatment and completed the FOTO knee FS PROM at admission and discharge.

Exclusions: •Patients who are not being treated for a Knee impairment

•<18 years of age

Adjustment/Stratification:

Level of Analysis: Facility, Clinician : Group/Practice, Clinician : Individual

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility, Other, Ambulatory Care : Outpatient Rehabilitation

Type of Measure: Outcome

Data Source: Patient Reported Data/Survey

Measure Steward: Focus on Therapeutic Outcomes, Inc

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

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

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

1a. Evidence: **Y-13**; **N-5; I-X**; 1b. Performance Gap: **H-0**; **M-7**; **L-5**; **I-7** 1c. High Priority: **H-7**; **M-11**; **L-1**; **I-0**

UPDATED VOTES FOR 1b. Performance Gap: H-4; **M-13**; **L-1**; **I-0**

Rationale:

- Please see discussion under measure 0423.

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

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

2a. Reliability: **H-2**; **M-4**; **L-7**; **I-6** 2b. Validity: **H-2**; **M-2**; **L-9**; **I-6 (informational vote only; non-binding)**

UPDATED VOTES FOR 2a. Reliability: H-4; **M-12**; **L-1**; **I-1 2b. Validity: H-3**; **M-14**; **L-1**; **I-0**

Rationale:

- Please see discussion under measure 0423.

3. Feasibility: H-7; M-8; L-3; I-1 (informational vote only; non-binding)

UPDATED VOTES FOR Feasibility: H-4; M-9; L-3; I-2

(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)

Rationale:

- Please see the discussion under measure 0423.

4. Use and Usability: H-4; M-5; L-7; I-3 (informational vote only; non-binding)

0422 Functional status change for patients with Knee impairments

UPDATED VOTES FOR Use and Usability: H-4; M-8; L-5; I-1

(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)

Rationale:

- Please see discussion under measure 0423.

5. Related and Competing Measures

- The Committee discussed whether this measure potentially competes with 2653: Average change in functional status following total knee replacement surgery (Minnesota Community Measurement). The Committee determined that the measures have different focus in terms of the target population, provider types, and clinical settings, as well as the clinical area. The developers indicated that the FOTO measure is broader and applicable to any kind of knee impairments, as opposed to measure 2653, which only focuses on patients with knee replacements. Therefore, the Committee agreed that the measures were related but not competing. The Committee did not make recommendations for harmonization.
- The Committee also considered the suite of FOTO measures to be related to 2624: Functional outcome assessment (CMS), as they address functional status for patients age 18 years and older. However, they differ in type as 2624 is a process measure, while FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related, but did not make recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-X**; **N-X; Y-17; N-1**

- Please see discussion under measure 0423.

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

Comments received:

- Comments on the set of FOTO measures (0422 – 0428) noted that the measure is stated to apply to patients age 14 and older. The Children's Hospital Association agrees with the Committee's discussion and request for evidence that the measure is understandable and appropriate for patients under the age of 18 as the measure was initially developed for patients 18 and over. Additional, comments stated that they are important patient centered outcomes, and while the measures are not perfect they could be improved as additional data is collected. Commenters also believed this measure should be considered for endorsement because it focuses on an important patient-centered outcome and addressees an important gap area for quality improvement. While this measure may not be perfect, it is an important patient centered outcome. The measure can be analyzed and improved as additional data is collected.

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:

- FOTO appreciates this support. The PFCC committee has requested and FOTO has provided additional analysis of validity and reliability at the clinician and clinic level. FOTO is committed to improvement of its measures and is involved in research to examine the relationship of its measures to other measures, including global ratings. In FOTO's survey development it has progressed from global measures to more body part specific measures because of the improved measure sensitivity realized with a specific vs. global measure, which FOTO believes is an important psychometric advantage. For example effect size was more than doubled comparing the FOTO Lumbar CAT (1.05) and AM PAC CAT Daily Activity Scale (.42) when applied to patients with spine impairments. This increased sensitivity has been noted in all of the body part data. There are also clinical advantages with the patient being presented with more pertinent (to

0422 Functional status change for patients with Knee impairments

their impairment) body part functional items to answer.

- The committee requested evidence that the instrument, which was originally developed for ages 18 and over, has been tested for understandability and appropriateness for youth down to age 14, as included in the measure. FOTO justified their initial request to change the inclusion criteria for its measures from 18 to 14 years old using the results of sensitivity analyses examining the impact of changing the age exclusion criteria on the risk adjustment models. However, in light of the discussions in committee, we recognize that additional testing is necessary. Therefore, we have requested permission to withdraw this change and return to the 18 years and older inclusion criteria. FOTO plans to perform studies on the understandability and appropriateness for youth 14-18 in the future. We have also recalculated the marginal means estimates by age-groups 18 and older.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0423 Functional status change for patients with Hip impairments

Submission |

Description: A self-report measure of change in functional status for patients 18 years+ with hip impairments. The change in functional status assessed using FOTO's (hip) PROM is adjusted to patient characteristics known to be associated with functional status outcomes (risk adjusted) and used as a performance measure at the patient level, at the individual clinician, and at the clinic level to assess quality.

Numerator Statement: Patient Level: The residual functional status score for the individual patient (residual scores are the actual change scores - predicted change after risk adjustment.

Individual Clinician Level: The average residuals in functional status scores in patients who were treated by a clinician in a 12 month time period for hip impairment.

Clinic Level: The average residuals in functional status scores in patients who were treated by a clinic in a 12 month time period for hip impairment.

Denominator Statement: All patients 18 years and older with hip impairments who have initiated rehabilitation treatment and complete the FOTO hip FS PROM at admission and discharge.

Exclusions: •Patients who are not being treated for a Hip impairment

•<18 years of age

Adjustment/Stratification:

Level of Analysis: Facility, Clinician : Group/Practice, Clinician : Individual

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility, Other, Ambulatory Care : Outpatient Rehabilitation

Type of Measure: Outcome

Data Source: Patient Reported Data/Survey

Measure Steward: Focus On Therapeutic Outcomes, Inc

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

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

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

1a. Evidence: **Y-13**; **N-5**; 1b. Performance Gap: **H-0**; **M-7**; **L-5**; **I-7** 1c. High Priority: **H-7**; **M-11**; **L-1**; **I-0**

UPDATED VOTES FOR 1b. Performance Gap: H-4; **M-13**; **L-1**; **I-0**

Rationale:

0423 Functional status change for patients with Hip impairments

- The Committee discussed the general question of whether functional status measures should include attributions to specific body parts. The advantage is that it can enhance the specificity of treatment, but it both limits changes in functional status to a particular body part and it also limits the degree to which clinics can be compared across different types of injuries.
- The developer stated that wrist and hand conditions would affect functional status much differently than foot or ankle conditions, for example, and that is why they have different forms. They also noted that clinicians wanted a more efficient tool that did not ask patients irrelevant items.
- The committee remained concerned that it is difficult to risk adjust away variability by the different types of injuries and diseases included, but did note approval of the items included in the types of function assessed. However, they pointed out that different types of injuries would cause very different abilities to heal; the developer thought the focus on functional outcomes would be similar in terms of what the patients care about.
- In response to questions, the developer clarified that the measures have short forms that are in the public domain. The short form is on paper and the full form is a CAT measure. The developer stated that the short form predicts 96-97% of the variance of the full measure so there is minimal bias introduced by the different forms. In addition, the actual CAT survey is publically posted and can be used via the developer's website, along with the coefficients from the risk models. They further explained that short form contains the most important items from the full bank, and that it has also been tested. The form has been calibrated to the original CAT.
- The committee agreed the information provided on the site, which includes the short form and the CAT, and the fact that a provider can use the tool, derive a score, and report it, all without subscribing, meets the criteria for publically available.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability</u>
(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)
2a. Reliability: **H-2**; **M-4**; **L-7**; **I-6** 2b. Validity: **H-2**; **M-2**; **L-9**; **I-6 (informational vote only; non-binding)**
UPDATED VOTES FOR 2a. Reliability: H-4; **M-12**; **L-1**; **I-1 2b. Validity: H-3**; **M-14**; **L-1**; **I-0**
<u>Rationale</u>:

- The Committee clarified what factors were risk adjusted for in the measure, and confirmed that the etiology of the hip impairment was not included in the risk adjustment. The developer agreed that it may be important in outcomes, but that it would be reflected in the intake functional assessment, and that they thought they could predict a fair amount of the variation in outcomes with characteristics they currently adjust for. Committee members agreed the best predictor of future function would be prior function for most functional status measures. The developers stated that a certain proportion of variability is also attributable to the clinic and the clinicians.
- The risk adjustment modeling includes all patients who have complete intake and discharge scores, but to calculate the performance measures, there are thresholds for participation. A committee member pointed out that the people who come in more often are more likely to get sampled and are more likely to get care, which could impact the link between process and outcome, and raised the concern that very small numbers could produce a lot of statistical noise. The developer explained that was why small sample sizes were excluded, and the committee member suggested hierarchical linear modeling to

0423 Functional status change for patients with Hip impairments
address this issue of floating sample sizes. • The committee requested more information about the interclass correlation coefficients at the clinician and clinic levels, and the developer offered to follow up with that information. The committee noted that this information would make sure that there is enough evidence that the measure is distinguishing clinicians from each other. • The developer noted that they had submitted supplementary information on the validity of the provider classification method, which showed that in the high-performing clinics, a greater proportion of patients improved more than a minimally clinically important amount. • Raising the concern that the risk adjustment for gender and payer may actually mask disparities in care, the committee requested more information and a justification for the risk adjustment variables, especially gender and payer. • They also requested evidence that the instrument, which was originally developed for ages 18 and over, has been tested for understandability and appropriateness for youth down to age 14, as included in the measure.
3. Feasibility: H-7; M-8; L-3; I-1 (informational vote only; non-binding) **UPDATED VOTES FOR Feasibility: H-4; M-9; L-3; I-2** *(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)* Rationale: • A committee member asked about the inclusion of fear or avoidance, stating that most are not collecting this information. The developer explained that they have found it is predictive of outcome, and that patients with higher levels of fear avoidance do not do as well, so they have incorporated it as a variable. • They explained that the time it takes to collect this information using the CAT is one to two minutes. The FOTO CAT and the Oswestry, which takes 6-8 minutes, are similar psychometrically but the FOTO tool is less of a burden to complete, record, and score. • The developer also noted that clinicians, patient managers, and patients were involved in the development of the items and that the CAT was well received by all three groups, for its efficiency and ease of use.
4. Use and Usability: H-4; M-5; L-7; I-3 (informational vote only; non-binding) **UPDATED VOTES FOR Use and Usability: H-4; M-8; L-5; I-1** *(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)* Rationale: • The committee had no additional comments regarding use and usability.
5. Related and Competing Measures • The Committee considered the suite of FOTO measures to be related to 2624: Functional Outcome Assessment (CMS), as they address functional status for patients age 18 years and older. However, they differ in type, as 2624 is a process measure, while FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related but did not make recommendations for harmonization.
Standing Committee Recommendation for Endorsement: Y-X; N-X; Y-17; N-1
6. Public and Member Comment: March 2, 2015- March 31, 2015

0423 Functional status change for patients with Hip impairments

- **Please see public and member comment under measure 0422.**

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0424 Functional status change for patients with Foot and Ankle impairments

Submission |

Description: A self-report measure of change in functional status for patients 18 years+ with foot and ankle impairments. The change in functional status assessed using FOTO's (foot and ankle) PROM is adjusted to patient characteristics known to be associated with functional status outcomes (risk adjusted) and used as a performance measure at the patient level, at the individual clinician, and at the clinic level to assess quality.

Numerator Statement: Patient Level: The residual functional status score for the individual patient (residual scores are the actual change scores - predicted change after risk adjustment)

Individual Clinician Level: The average of residuals in functional status scores in patients who were treated by a clinician in a 12 month time period for foot and or ankle impairment.

Clinic Level: The average of residuals in patients who were treated by a clinic in a 12 month time period for foot and or ankle impairment.

Denominator Statement: All patients 18 years and older with foot or ankle impairments who have initiated rehabilitation treatment and completed the FOTO foot and ankle PROM at admission and discharge

Exclusions: •Patients who are not being treated for a foot and ankle impairment

•<18 years of age

Adjustment/Stratification:

Level of Analysis: Facility, Clinician : Group/Practice, Clinician : Individual

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility, Other, Ambulatory Care : Outpatient Rehabilitation

Type of Measure: Outcome

Data Source: Paper Medical Records

Measure Steward: Focus on Therapeutic Outcomes, Inc

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

1. Importance to Measure and Report: The measuremeet the Importance criteria

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

1a. Evidence: **Y-13**; **N-5; I-X**; 1b. Performance Gap: **H-0**; **M-7**; **L-5**; **I-7** 1c. High Priority: **H-7**; **M-11**; **L-1**; **I-0**

UPDATED VOTES FOR 1b. Performance Gap: H-4; **M-13**; **L-1**; **I-0**

Rationale:

- Please see discussion under measure 0423.

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

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

2a. Reliability: **H-2**; **M-4**; **L-7**; **I-6** 2b. Validity: **H-2**; **M-2**; **L-9**; **I-6 (informational vote only; non-binding)**

UPDATED VOTES FOR 2a. Reliability: H-4; **M-12**; **L-1**; **I-1 2b. Validity: H-3**; **M-14**; **L-1**; **I-0**

Rationale:

0424 Functional status change for patients with Foot and Ankle impairments

- Please see discussion under measure 0423.

3. Feasibility: H-7; M-8; L-3; I-1 (informational vote only; non-binding)

UPDATED VOTES FOR Feasibility: H-4; M-9; L-3; I-2

(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)

Rationale:

- Please see discussion under measure 0423.

4. Use and Usability: H-4; M-5; L-7; I-3 (informational vote only; non-binding)

UPDATED VOTES FOR Use and Usability: H-4; M-8; L-5; I-1

(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)

Rationale:

- Please see discussion under measure 0423.

5. Related and Competing Measures

- The Committee considered the suite of FOTO measures to be related to 2624: Functional Outcome Assessment (CMS), as they address functional status for patients age 18 years and older. However, they differ in type, as 2624 is a process measure, while FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related but did not make recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-X**; **N-X; Y-17; N-1**

- Please see discussion under measure 0423.

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

- **Please see public and member comment under measure 0422.**

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0425 Functional status change for patients with lumbar impairments

Submission |

Description: A self-report outcome measure of functional status for patients 18 years+ with lumbar impairments. The change in functional status assessed using FOTO (lumbar) PROM is adjusted to patient characteristics known to be associated with functional status outcomes (risk adjusted) and used as a performance measure at the patient level, at the individual clinician, and at the clinic level by to assess quality.

Numerator Statement: Patient Level: The residual functional status score for the individual patient (residual scores are the actual change scores - predicted change after risk adjustment).

Individual Clinician Level: The average of residuals in functional status scores in patients who were treated by a clinician in a 12 month time period for lumbar impairment.

Clinic Level: The average of residuals) in functional status scores in patients who were treated by a clinic in a 12 month time period for lumbar impairment.

Denominator Statement: All patients 18 years and older with a lumbar impairment who have initiated rehabilitation treatment and completed the FOTO (lumbar) PROM.

Exclusions: •Patients who are not being treated for a lumbar impairment

0425 Functional status change for patients with lumbar impairments
•<18 years of age **Adjustment/Stratification**: **Level of Analysis:** Facility, Clinician : Group/Practice, Clinician : Individual **Setting of Care:** Ambulatory Care : Clinician Office/Clinic, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility, Other, Ambulatory Care : Outpatient Rehabilitation **Type of Measure**: Outcome **Data Source**: Patient Reported Data/Survey **Measure Steward**: Focus on Therapeutic Outcomes, Inc
STANDING COMMITTEE MEETING [01/21/2015-01/22/2015] **1. Importance to Measure and Report: The measuremeets the Importance criteria** (1a. Evidence: 1b. Performance Gap, 1c. High Priority) 1a. Evidence: **Y-13**; **N-5; I-X**; 1b. Performance Gap: **H-0**; **M-7**; **L-5**; **I-7** 1c. High Priority: **H-7**; **M-11**; **L-1**; **I-0** **UPDATED VOTES FOR 1b. Performance Gap: H-4**; **M-13**; **L-1**; **I-0** Rationale: • Please see discussion under measure 0423.
2. Scientific Acceptability of Measure Properties: The measure meets the Scientific Acceptability (2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity) 2a. Reliability: **H-2**; **M-4**; **L-7**; **I-6** 2b. Validity: **H-2**; **M-2**; **L-9**; **I-6 (informational vote only; non-binding)** **UPDATED VOTES FOR 2a. Reliability: H-4**; **M-12**; **L-1**; **I-1 2b. Validity: H-3**; **M-14**; **L-1**; **I-0** Rationale: • Please see discussion under measure 0423.
3. Feasibility: H-7; M-8; L-3; I-1 (informational vote only; non-binding) **UPDATED VOTES FOR Feasibility: H-4; M-9; L-3; I-2** *(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)* Rationale: • Please see discussion under measure 0423.
4. Use and Usability: H-4; M-5; L-7; I-3 (informational vote only; non-binding) **UPDATED VOTES FOR Use and Usability: H-4; M-8; L-5; I-1** *(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)* Rationale: • Please see discussion under measure 0423.
5. Related and Competing Measures • The Committee considered the suite of measures to be related to 2624: Functional Outcome Assessment (CMS), as they address functional status for patients age 18 years and older. However, they differ in type, as 2624 is a process measure, while FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related but did not make recommendations for harmonization.
Standing Committee Recommendation for Endorsement: **Y-X**; **N-X; Y-17; N-1** • Please see discussion under measure 0423.
6. Public and Member Comment: March 2, 2015- March 31, 2015 • **Please see public and member comment under measure 0422.**
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

0425 Functional status change for patients with lumbar impairments

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0426 Functional status change for patients with Shoulder impairments

Submission |

Description: A self-report outcome measure of change in functional status for patients 18 years+ with shoulder impairments. The change in functional status assess using FOTO's (shoulder) PROM is adjusted to patient characteristics known to be associated with functional status outcomes (risk adjusted) and used as a performance measure at the patient level, at the individual clinician, and at the clinic level to assess quality.

Numerator Statement: Patient Level: The residual functional status score for the individual patient (residual scores are the actual change scores - predicted change after risk adjustment.

Individual Clinician Level: The average of residuals in functional status scores in patients who were treated by a clinician in a 12 month time period for shoulder impairment.

Clinic Level: The average of residuals in functional status scores in patients who were treated by a clinic in a 12 month time period for shoulder impairment.

Denominator Statement: All patients 18 years and older with shoulder impairments who have initiated rehabilitation treatment and completed the FOTO shoulder FS outcome instrument at admission and discharge.

Exclusions: •Patients who are not being treated for a Shoulder impairment

•<18 years of age

Adjustment/Stratification:

Level of Analysis: Facility, Clinician : Group/Practice, Clinician : Individual

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Home Health, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility, Other, Ambulatory Care : Outpatient Rehabilitation

Type of Measure: Outcome

Data Source: Patient Reported Data/Survey

Measure Steward: Focus on Therapeutic Outcomes, Inc

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 Priority)

1a. Evidence: **Y-13**; **N-5; I-X**; 1b. Performance Gap: **H-0**; **M-7**; **L-5**; **I-7** 1c. High Priority: **H-7**; **M-11**; **L-1**; **I-0**

UPDATED VOTES FOR 1b. Performance Gap: H-4; **M-13**; **L-1**; **I-0**

Rationale:

- Please see discussion under measure 0423.

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

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

2a. Reliability: **H-2**; **M-4**; **L-7**; **I-6** 2b. Validity: **H-2**; **M-2**; **L-9**; **I-6 (informational vote only; non-binding)**

UPDATED VOTES FOR 2a. Reliability: H-4; **M-12**; **L-1**; **I-1 2b. Validity: H-3**; **M-14**; **L-1**; **I-0**

Rationale:

- Please see discussion under measure 0423.

3. Feasibility: H-7; M-8; L-3; I-1 (informational vote only; non-binding)

0426 Functional status change for patients with Shoulder impairments

UPDATED VOTES FOR Feasibility: H-4; M-9; L-3; I-2

(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)

Rationale:

- Please see discussion under measure 0423.

4. Use and Usability: H-4; M-5; L-7; I-3 (informational vote only; non-binding)

UPDATED VOTES FOR Use and Usability: H-4; M-8; L-5; I-1

(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)

Rationale:

- Please see discussion under measure 0423.

5. Related and Competing Measures

- The Committee considered this measure to be related to 2624: Functional Outcome Assessment (CMS), as they address functional status for patients age 18 years and older. However, they differ in type, as 2624 is a process measure, while FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related, but did not make recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-X**; **N-X; Y-17; N-1**

- Please see discussion under measure 0423.

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

- **Please see public and member comment under measure 0422.**

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

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

Submission |

Description: A self-report outcome measure of functional status for patients 18 years+ with elbow, wrist, hand impairments. The change in functional status assessed using FOTO (elbow, wrist and hand) PROM is adjusted to patient characteristics known to be associated with functional status outcomes (risk adjusted) and used as a performance measure at the patient level, at the individual clinician, and at the clinic level to assess quality.

Numerator Statement: Patient Level: The residual functional status score for the individual patient (residual scores are the actual change scores - predicted change after risk adjustment).

Individual Clinician Level: The average of residuals in functional status scores in patients who were treated by a clinician in a 12 month time period for elbow, wrist and hand impairment.

Clinic Level: The average of residuals in functional status scores in patients who were treated by a clinic in a 12 month time period for elbow, wrist and hand impairments.

Denominator Statement: All patients 18 years and older with elbow, wrist or hand impairments who have initiated rehabilitation treatment and completed the FOTO (elbow, wrist and hand) PROM.

Exclusions: •Patients who are not being treated for an elbow, wrist and/or hand impairment

•<18 years of age

Adjustment/Stratification:

Level of Analysis: Facility, Clinician : Group/Practice, Clinician : Individual

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

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility, Other, Ambulatory Care : Outpatient Rehabilitation

Type of Measure: Outcome

Data Source: Patient Reported Data/Survey

Measure Steward: Focus on Therapeutic Outcomes, Inc

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 Priority)

1a. Evidence: **Y-13**; **N-5; I-X**; 1b. Performance Gap: **H-0**; **M-7**; **L-5**; **I-7** 1c. High Priority: **H-7**; **M-11**; **L-1**; **I-0**

UPDATED VOTES FOR 1b. Performance Gap: H-4; **M-13**; **L-1**; **I-0**

<u>Rationale</u>:

- Please see discussion under measure 0423.

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

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

2a. Reliability: **H-2**; **M-4**; **L-7**; **I-6** 2b. Validity: **H-2**; **M-2**; **L-9**; **I-6 (informational vote only; non-binding)**

UPDATED VOTES FOR 2a. Reliability: H-4; **M-12**; **L-1**; **I-1 2b. Validity: H-3**; **M-14**; **L-1**; **I-0**

<u>Rationale</u>:

- Please see discussion under measure 0423.

3. Feasibility: H-7; M-8; L-3; I-1 (informational vote only; non-binding)

UPDATED VOTES FOR Feasibility: H-4; M-9; L-3; I-2

(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)

<u>Rationale</u>:

- Please see discussion under measure 0423.

4. Use and Usability: H-4; M-5; L-7; I-3 (informational vote only; non-binding)

UPDATED VOTES FOR Use and Usability: H-4; M-8; L-5; I-1

(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)

<u>Rationale</u>:

- Please see discussion under measure 0423.

5. Related and Competing Measures

- The Committee considered the suite of FOTO measures to be related to 2624: Functional Outcome Assessment (CMS), as they address functional status for patients age 18 years and older. However, they differ in type, as 2624 is a process measure, while the FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related but did not make recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-X**; **N-X; Y-17; N-1**

- Please see discussion under measure 0423.

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

- Please see public and member comment under measure 0422.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0428 Functional status change for patients with General orthopaedic impairments

Submission |

Description: A self-report outcome measure of functional status for patients 18 years+ with general orthopaedic impairments. The change in functional status assessed using FOTO (general orthopedic) PROM is adjusted to patient characteristics known to be associated with functional status outcomes (risk adjusted) and used as a performance measure at the patient level, at the individual clinician, and at the clinic level by to assess quality.

Numerator Statement: Patient Level: The residual functional status score for the individual patient (residual scores are the actual change scores - predicted change after risk adjustment).

Individual Clinician Level: The average of residuals in functional status scores in patients who were treated by a clinician in a 12 month time period for general orthopaedic impairment.

Clinic Level: The average of residuals in functional status scores in patients who were treated by a clinic in a 12 month time period for general orthopaedic impairment.

Denominator Statement: All patients 18 years and older with general orthopaedic impairments who have initiated rehabilitation treatment and completed the FOTO (general orthopaedic) PROM.

Exclusions: •Patients who are not being treated for a General orthopaedic impairment

•<18 years of age

Adjustment/Stratification:

Level of Analysis: Facility, Clinician : Group/Practice, Clinician : Individual

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility, Other, Ambulatory Care : Outpatient Rehabilitation

Type of Measure: Outcome

Data Source: Paper Medical Records

Measure Steward: Focus on Therapeutic Outcomes, Inc

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 Priority)

1a. Evidence: **Y-13**; **N-5** 1b. Performance Gap: **H-0**; **M-7**; **L-5**; **I-7** 1c. High Priority: **H-7**; **M-11**; **L-1**; **I-0**

UPDATED VOTES FOR 1b. Performance Gap: H-4; **M-13**; **L-1**; **I-0**

<u>Rationale</u>:

- Since there are a number of body-part specific measures, the committee requested the rationale behind the general measure. The developer explained that this measure looks at impairments of the cervical and thoracic vertebrae, the ribs, TMJ, etc. They further explained that this data is the predominant impairment group, covering 70% of the data, but that there were 30-40% of the data that covered other body parts, so they elected to keep it as a general measure rather than a measure focused on the cervical vertebrae. They noted that they are developing a measure specifically focused on cervical but that it is not yet completed.
- Please see remaining discussion under measure 0423.

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-2**; **M-4**; **L-7**; **I-6** 2b. Validity: **H-2**; **M-2**; **L-9**; **I-6 (informational vote only; non-binding)**

UPDATED VOTES FOR 2a. Reliability: H-4; **M-12**; **L-1**; **I-1 2b. Validity: H-3**; **M-14**; **L-1**; **I-0**

0428 Functional status change for patients with General orthopaedic impairments

Rationale:

- Please see discussion under measure 0423.

3. Feasibility: H-7; M-8; L-3; I-1 (informational vote only; non-binding)

UPDATED VOTES FOR Feasibility: H-4; M-9; L-3; I-2

(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)

Rationale:

- Please see discussion under measure 0423.

4. Use and Usability: H-4; M-5; L-7; I-3 (informational vote only; non-binding)

UPDATED VOTES FOR Use and Usability: H-4; M-8; L-5; I-1

(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)

Rationale:

- Committee members noted there is a logic to the body-part specific surveys, but also noted that the different instruments have a lot of similarities and wondered if it really makes sense to have so many surveys, given the cost and time to implement different surveys for each patient, or if a more holistic model would be appropriate. They also mentioned that patients say that they are whole people, not just body parts. The developer agreed there are similarities, but explained that when they analyzed the data, they found that difficultly levels for various items change depending on the body part affected. The measure is more precise if they recalibrate the difficulty level for each item for each body part. They also agreed that providers need to address the total body to improve function, and should be doing so at each visit.

5. Related and Competing Measures

- The Committee considered the suite of FOTO measures to be related to 2624: Functional Outcome Assessment (CMS), as they address functional status for patients age 18 years and older. However, they differ in type, as 2624 is a process measure, while the FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related but did not make recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-X**; **N-X; Y-17; N-1**

- Please see discussion under measure 0423.

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

- Please see public and member comment under measure 0422.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

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

Submission |

Description: This measure, based on data from the Minimum Data Set (MDS) 3.0 assessment of long-stay nursing facility residents, estimates the percentage of long-stay residents in a nursing facility whose need for assistance

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

with late-loss Activities of Daily Living (ADLs), as reported in the target assessment, increased when compared with a prior assessment. The four late-loss ADLs are: bed mobility, transfer, eating, and toilet use. This measure is calculated by comparing the change in each ADL item between the target assessment (OBRA, PPS or discharge) and a prior assessment (OBRA, PPS or discharge). Long-stay nursing facility residents are those with a nursing facility stay of 101 cumulative days or more.

Numerator Statement: The numerator is the number of long-stay residents who have a selected target MDS assessment (OBRA, PPS, or discharge) reporting a defined amount of decline in ADL function when compared with a prior assessment (OBRA, PPS, or discharge). This decline in function is captured as an increase in the resident's need for assistance with late-loss ADLs, when compared with the resident's prior assessment, indicated by a higher score on the applicable MDS items on the more recent assessment (which are coded such that a higher score indicates the need for more assistance with an ADL task). Late-loss ADL items are bed mobility, transfer, eating, and toilet use. The threshold increase in need for assistance (suggesting decline in function) that results in a resident being counted in the numerator is met if the score for at least one late-loss ADL item increases by two or more points or if the score for two or more of the late-loss ADLs items increase by one point. The typical interval between the target and prior assessment dates is approximately 90 days.

Denominator Statement: The denominator includes all long-stay residents with a selected target MDS assessment (OBRA, PPS, or discharge) during the quarter and a prior assessment who did not meet the exclusion criteria. Long-stay residents are defined as residents who have stayed in the nursing home for 101 cumulative days or more.

Exclusions: There are six exclusions applied to the denominator: (1) self-performance total dependence on all four late-loss ADL items during the prior assessment (and therefore it is not possible for the resident to decline sufficiently to be counted in the numerator), (2) self-performance total dependence on three late-loss ADL items during the prior assessment and self-performance extensive assistance on the fourth late-loss ADL item (and therefore it is not possible for the resident to decline sufficiently to be counted in the numerator), (3) comatose status on the target assessment, (4) prognosis of life expectancy of less than six months on the target assessment, (5) receiving hospice care on the target assessment, or/and (6) the resident is not in the numerator and has missing values for any of the four ADL items on the target or prior assessment.

Nursing facilities are excluded from public reporting if their denominator size is less than 30 residents.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility

Type of Measure: Outcome

Data Source: Electronic Clinical Data

Measure Steward: Centers for Medicare & Medicaid Services

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-17**; **N-1**; 1b. Performance Gap: **H-8**; **M-8**; **L-0**; **I-2;** 1c. High Priority: **H-11**; **M-6**; **L-1**; **I-0**

<u>Rationale</u>:

- The Committee agreed that the therapeutic goal to delay decline in the selected ADLs is very important for this population, but raised concerns about the exclusions in the denominator. Of particular concern was the exclusion for people with less than six months expected survival, which could have potential risk for gaming, as well as the difficulty in reliably identifying people with less than six months expected survival.
- The measure developers explained that this exclusion has multiple intentions. One is that if people are at

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

end of life, they are going to be at much higher risk for ADL decline. Second, there may be unintended consequences for patients in hospice care; facilities may not be willing to stop providing interventions intended to maintain function, despite a patient's end of life preferences.

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-3**; **M-11**; **L-2**; **I-1** 2b. Validity: **H-0**; **M-12**; **L-5**; **I-1**

<u>Rationale</u>:

- In general, the Committee noted that reliability was good for this measure. One Committee member agreed that reliability at the patient level was well defined, but expressed concerns regarding both reliability and validity testing at the facility level.
- The developers noted that on Nursing Home Compare, they are looking at multiple averages of weighted data at a facility-level, across multiple calendar quarters, to provide the average for the entire country per state. Additionally, the developers stated that this measure only reports on basic outcomes, but in the future it could be revised or paired with other measures to demonstrate many more outcomes.

4. Feasibility: H-12; M-6; L-0; 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 noted that the data elements are discrete and electronically captured.

3. Use and Usability: H-10; M-7; L-0; I-1

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

<u>Rationale</u>:

- The Committee acknowledged that the measure is publicly reported on Nursing Home Compare and used in benchmarking.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-15**; **N-3**

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

Comments received:

- This measure received two comments indicating a lack of support for endorsement. The rationale focused on the fact that many SNF patients are working on maintaining function, not improvement, and that improvement should be happening before patients are moved to SNFs.
- Another commenter did not specify support or lack of support, but raised two concerns with this measure, first, the potential disincentives for maintenance therapy, and second, the related concern that the list of exclusions is not broad enough.

Committee response:

- During the in-person meeting, the Committee raised similar concerns about this measure, but ultimately agreed that the therapeutic goal to delay decline in the selected ADLs is very important for this population. While the Committee raised concerns about the exclusions in the denominator, the discussion was mainly about the reliability of identifying people with less than six months expected survival. The measure developers explained that there are multiple intentions with regard to this exclusion. One is that

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

if people are at end of life, they are going to be at much higher risk for ADL decline. On the other hand, if they are included in the measure there may be an unintended consequence where facilities may not be willing to set aside some interventions that they need to do in order to maintain function and thus not respecting preferences at end of life.

Developer response:

- NQF #0688 tracks potential decline in function by measuring "the percent of residents whose need for help with activities of daily living (ADL) has increased." Accordingly, the purpose of this measure is to assess decline in ADL function among long-stay nursing home residents. This change in ADL function is documented during the period of nursing home stay by comparing ADL function from one nursing home assessment to the next. We agree that the goal of many long-stay residents is to maintain their existing ADLs and may not be focused on ADL improvement; we believe that NQF #0688 is aligned with this perspective, as it is not focused on improvement. A higher score for this measure indicates lower quality. Patients maintaining their level of functional ability for the 4 late-loss ADLs would NOT be counted in the numerator for this measure and would be considered as experiencing good quality. We also believe that NQF #0688 is not at odds with other potential measures described by the commenter that would focus on improving ADLs in other settings prior to nursing home admission. However, the measure proposed by the commenter might be more appropriate for short-stay nursing home residents who are generally admitted for goals different from long-stay residents.
- NQF #0688 is an outcome measure defined as "the percent of residents whose need for help with activities of daily living (ADL) has increased." Accordingly, the purpose of this measure is to assess decline in ADL function among long-stay nursing home residents, rather than improvement. We agree that the goal for many long-stay residents is trying to maintain their level of activity, thus focusing on maintenance, not improvement. We believe that the focus of NQF #0688 is aligned with this perspective by quantifying the proportion of long-stay residents who have experienced a loss in function. Residents are counted in the numerator of this measure if they experience an increase in need for assistance with late-loss ADLs in a given assessment period, as compared to a prior assessment. A higher score for this measure indicates lower quality. Thus patients maintaining functional ability for the 4 late-loss ADLs would NOT be counted in the numerator for this measure and would be considered as experiencing good quality.
- This measure (NQF #0688), the percent of residents whose need for help with activities of daily living (ADL) has increased (long stay), is designed to track decline in ADL function among long-stay nursing home residents from one assessment period to the next. CMS understands that improvement and recovery are not always feasible among long-stay nursing home resident populations, hence the appropriateness of using this measure to monitor increased needs for assistance (i.e., functional decline), rather than improvement for the long-stay nursing home resident population. The measure is designed so that each instance of a resident maintaining functional status is counted as an indicator of good facility quality. This comment references the Standing Committee recommendation to add exclusions to this measure, but these recommended exclusions noted in the Draft Report for Comment apply to measures of ADL improvement, whereas NQF #0688 measures ADL decline. As it stands, this measure has four exclusion groups: currently comatose, prognosis of life expectancy less than 6 months, receiving hospice care, or total dependence for all four ADL items on prior assessment. The Standing Committee suggested that there may be a potential for gaming, particularly with the six month prognosis item. We suggest that the item used to identify residents who have a prognosis of less than six months to live has relatively little risk for gaming because it is based on physician documentation in the medical record, rather than the clinical judgment of facility staff completing the assessment. In addition this exclusion applies to only a small number of residents, and the proportion of residents excluded from the measure for this reason has declined over time, which does not support the suggestion that it is an exclusion that is being gamed (1).

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

While there is concern regarding physicians' ability to identify end of life prognosis, analyses of residents included in this measure (i.e., greater than six month prognosis) show that very few (3.3%) expired. In addition, item level reliabilities were very high when tested during the RAND development of the MDS 3.0 (gold standard to gold standard nurse kappa: 0.872; gold standard nurse to facility nurse kappa: 0.964) (2). Lastly, we reiterate that including end of life residents in the measure could not only put them at risk for reduced access, but also at risk for care at odds with end-of-life goals and patient preferences. With regard to the commenter's concerns that other high risk populations should be added to the exclusions, we will continue to analyze and monitor this measure for conditions that should be excluded.

(1) RTI analyses of MDS 3.0 data show that in Quarter 1 of 2011, 0.4% of long-stay residents were excluded due to less than six-month prognosis. This proportion declined to 0.08% in quarter 2 of 2012, rebounded slightly to 0.13% in Quarter 4 of 2012, and declined again to 0.11% in Quarter 2 of 2013 where it has held steady at 0.11% through Quarter 2 of 2014. (RTI programming reference: nh_22_10, all quarters through 13_14).

(2)Saliba D., Buchanan D., Development and Validation of a Revised Nursing Home Assessment Tool: MDS 3.0 Appendices. Prepared for the Centers for Medicare & Medicaid Services. April 2008. Available at: http://www.geronet.med.ucla.edu/centers/borun/Appendix_A-G.pdf

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0701 Functional Capacity in COPD patients before and after Pulmonary Rehabilitation

Submission |

Description: The percentage of patients with COPD who are found to increase their functional capacity by at least 25 meters (82 feet), as measured by a standardized 6 minute walk test (6MWT) after par-ticipating in pulmonary rehabilitation (PR).

Numerator Statement: Number of patients who are found to increase their functional capacity by at least 25 meters (82 feet), as measured by 6MWT distance at PR program entry and completion.

Denominator Statement: All patients with clinician diagnosed COPD at PR program entry who completed PR during the measurement period and who completed at least 10 PR sessions within 3 months of PR program entry.

Exclusions: Patients for whom a 6MWT would be contraindicated due to acute or unstable medical conditions

Patients who are unable to perform a 6MWT due to orthopedic, neurological, cognitive or psychiatric impairments and/or safety reasons.

Patients who have not completed at least 10 PR sessions within 3 months of program entry.

Adjustment/Stratification:

Level of Analysis: Facility, Clinician : Group/Practice, Clinician : Individual

Setting of Care: Ambulatory Care : Outpatient Rehabilitation

Type of Measure: Outcome

Data Source: Management Data, Electronic Clinical Data : Registry

Measure Steward: American Association of Cardiovascular and Pulmonary Rehabilitation

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)

0701 Functional Capacity in COPD patients before and after Pulmonary Rehabilitation

1a. Evidence: **Y-15**; **N-2**; 1b. Performance Gap: **H-4**; **M-8**; **L-0**; **I-5;** 1c. High Priority: **H-7**; **M-10**; **L-0**; **I-0**

Rationale:

- The Committee noted that there is a clear link between exercise training and its value in improving COPD function but expressed some concerns that that data only exist at the patient level, with no quality information at the program or facility level to compare differences or to identify how reliable or reproducible these scores are at the facility or program level. The lack of facility or program data also impacts the ability to assess performance gap.
- NQF stated that since this is an outcome measure, as opposed to a patient reported outcome measure, it would meet the criteria by providing the patient level result or the measure level result, which has been provided by the developer. The developer reiterated that the measurement is at the patient level and it is for programs to measure their changes in functional capacity. In the future, they plan to have programs compare to benchmarks.

2. Scientific Acceptability of Measure Properties:

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

2a. Reliability: **H-3**; **M-6**; **L-3**; **I-4** (consensus not reached) 2b. Validity: **H-3**; **M-10**; **L-3**; **I-1 (consensus reached)**

UPDATED VOTES FOR 2a. Reliability: H-4; M-15; L-0; I-0

Rationale:

- The committee once more raised concerns about not having any evidence beyond the patient level data. NQF stated that ideally data for both levels are preferred, but is not a requirement to endorse. The developer noted that a six minute walk test is done based on guidelines using a standardized tool, which is one of the reasons why the evidence from the previous literature can be used to show the measure is valid and reliable.
- One Committee member recommended that the measure should demonstrate a percent improvement in performance, as opposed to a specific number, since it could be clinically reasonable to migrate from a six minute walk test to a two minute test or other performance tests that are shorter. The developer responded that this would need a careful data analysis from the pulmonary rehab database to establish these cut points and this would change from the evidence base currently practiced, but they would certainly consider this suggestion in the future.

4. Feasibility: H-6; M-11; L-0; 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 noted that data would be generated routinely and can be electronically submitted and abstracted. According to the developer, the measure can be submitted to the AACVPR Outpatient Pulmonary Rehabilitation Registry or another data base for quality improvement on a standardized data collection form, as recommended in the American Thoracic Society (ATS) guidelines for administration of the six minute walk test. The guidelines for administration are provided to all programs in the AACVPR PR Outcomes Resource Guide as well as published in ATS guidelines.

3. Use and Usability: H-7; M-9; L-1; I-0

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

Rationale:

0701 Functional Capacity in COPD patients before and after Pulmonary Rehabilitation

- This measure is currently in use for quality improvement (internal to the specific organization). The developer plans on using this measure for public reporting.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-1; Y-19; N-0**

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

Comments received:

- One commenter requested a clarification in the specifications regarding ages.

Developer response:

- The age range is greater than or equal to 18 years old, with no upper limit.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

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

2286 Functional Change: Change in Self Care Score

(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: 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 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

2286 Functional Change: Change in Self Care Score
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 is superior. Therefore, these measures will both continue on to NQF member vote as recommended and be brought forward to the CSAC on June 9th for resolution of the best-in-class decision. • 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

2286 Functional Change: Change in Self Care Score

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. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2287 Functional Change: Change in Motor Score

Submission |

Description: Change in rasch derived values of motor 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 12 FIM® items:Feeding, Grooming, Dressing Upper Body, Dressing Lower Body, Toileting, Bowel, Expression, Memory, Transfer Bed/Chair/Wheelchair, Transfer Toilet, Locomotion and Stairs.

Numerator Statement: Average change in rasch derived motor functional score from admission to discharge at the facility level. 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 IRF or patients who died within the IRF 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: 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: Administrative claims, Other

Measure Steward: Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities, Inc. and

2287 Functional Change: Change in Motor Score

its successor in interest, UDSMR, LLC.

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-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**

<u>Rationale</u>:

- This is one of a suite of measures derived from the FIM; Measure 2286 calculates and reports a change in self-care score; measure 2321 reports a change in mobility score, and together they comprise this measure, 2287, which calculates a change in motor score. The developer indicated it was important that the committee 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'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 had been 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 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

2287 Functional Change: Change in Motor Score
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 the set of improvements in ADLs for home health measures, including 0167, 0174, and 0175, as these measures have the same focus area, but are specified for different settings and populations. Measures 0167, 0174, and 0175 are intended for certified home health patients ages 18 and above, and measure 2287 was submitted as an inpatient rehabilitation facility measure. The Committee agreed that there was a need for all of the aforementioned measures and thus made no recommendations for harmonization. In alignment with the IMPACT Act provisions, the Committee emphasized the importance of using cross-setting measures in programs for future considerations.
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.

2287 Functional Change: Change in Motor Score

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. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

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)

2321 Functional Change: Change in Mobility Score

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: <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**

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

2321 Functional Change: Change in Mobility Score
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 is superior. Therefore, these measures will both continue on to NQF member vote as recommended and be brought forward to the CSAC on June 9th for resolution of the best-in-class decision.
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.

2321 Functional Change: Change in Mobility Score

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. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2612 CARE: Improvement in Mobility

Submission |

Description: The measure calculates a skilled nursing facility's (SNFs) average change in mobility for patients admitted from a hospital who are receiving therapy. The measure calculates the average change in mobility score between admission and discharge for all residents admitted to a SNF from a hospital or another post-acute care setting for therapy (i.e., PT or OT) regardless of payor status. This is a risk adjusted outcome measure, based on the mobility subscale of the Continuity Assessment and Record Evaluation (CARE) Tool and information from the admission MDS 3.0 assessment. The measure is calculated on a rolling 12 month, average updated quarterly.

Numerator Statement: The measure assesses the change in mobility. The numerator is the risk adjusted sum of the change in the CARE Tool mobility subscale items between admission and discharge for each individual admitted from a hospital or another post acute care setting regardless of payor status and are receiving therapy (PT or OT) for any reason in a skilled nursing center.

Denominator Statement: The denominator includes all residents admitted to a SNF from a hospital or another post-acute care setting who receive either PT or OT therapy for any reason during their stay regardless of payor status, have a completed mobility CARE tool assessment at admission and discharge and do not meet any of the exclusion criteria. The mobility items used from the CARE tool are listed below and rated on a 1-6 scale (see Appendix for copy of the mobility CARE tool assessment).

The items included in the CARE Tool Mobility subscale include:

- B1. Lying to Sitting on Side of Bed
- B2. Sit to Stand
- B3. Chair/Bed to Chair Transfer
- B4. Toilet Transfer
- B5a & B5b. Walking or Wheelchair Mobility
- C3. Roll left / right
- C4. Sit to Lying
- C5. Picking up object
- C7a. One Step Curb

2612 CARE: Improvement in Mobility

- C7b. Walk 50 ft. with Two Turns
- C7c. Walk 12 Steps.
- C7d. Walk Four Steps
- C7e. Walking 10 ft. on Uneven Surface
- C7f. Car Transfer

Exclusions: Patients are excluded for two broad reasons:

1. if they have conditions where improvement in mobility is very unlikely,

OR

2. have missing data necessary to calculate the measure

Additionally, facilities with denominator size of fewer than 30 patients during a 12 month period are excluded from reporting their data.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility

Type of Measure: Outcome

Data Source: Electronic Clinical Data, Other

Measure Steward: American Health Care Association

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-5**; **M-9**; **L-1**; **I-1;** 1c. High Priority: **H-10**; **M-5**; **L-1**; **I-0**

<u>Rationale</u>:

- The Committee agreed that the rationale supports a health outcome (change of mobility) in relation to the intervention of therapeutic services which are provided within the SNF. However, the Committee was unclear as to what a meaningful change in function would actually be for these patients, and how the measure as proposed relates to quality of care and patient outcomes related to returning home (i.e., basic mobility skills, ADLs and IADLs activities).
- The Committee expressed major concerns about the measure's focus on "improvement" in mobility when the Jimmo v. Sebelius settlement prevents requiring improvement in function as a condition of coverage in SNFs (as well as home health and outpatient services). Therefore, without appropriate risk adjustment, a SNF may be more likely to deny access to patients who require SNF services to maintain or prevent further deterioration of function.
- The Committee noted that there were significant variation and room for improvement in terms of performance gap.
- The Committee further noted that specific data on disparities was not included in this measure as specified by current NQF requirements. The Committee was interested in the inclusion of questions related to cognitive function. The developer stated that they risk-adjust for cognitive status using the MDS, however, if the CARE tool was inserted in the MDS, OASIS, and IRF-PAI they would be better able to collect patients' overall cognitive status.

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)

2612 CARE: Improvement in Mobility

2a. Reliability: **H-7**; **M-6**; **L-1**; **I-2** 2b. Validity: **H-5**; **M-10**; **L-0**; **I-1**

Rationale:

- The Committee determined that the measure specifications were precise, noting that the specifications were consistent with the evidence presented.
- Empiric reliability testing was performed at the patient level using data from the CARE tool. The Committee noted that reliability at the facility level was not provided and that it would be beneficial for the Committee to see the inter-classical correlation coefficients, a thumb print that suggests that there is a reproductive score within facilities that can then be used to compare between facilities.
- Empiric validity testing was conducted comparing the mobility measure set to other SNF quality measures, including the 5 star rating, Nursing Home Compare, and some specific measures such as pressure ulcers and rehospitalization. There was no correlation between the mobility measure and the 5 star rating, but there were variable correlations between the mobility measure and Nursing Home Compare. The Committee was specifically interested in the 30 day SNF risk adjusted rehospitalization rates, where the developer's hypothesis (i.e., improved mobility leads to lower rehospitalization) did not come to fruition. Instead, validity testing showed that higher rehospitalization rates were positively correlated with a higher mobility score.
- The Committee supported the developer's decision to exclude specific patients (e.g., ventilator patients, persistent coma, quadriplegic, hospice, and children) from the measure, given certain concerns regarding failure to improve, as well as other unintended consequences associated with treating certain high-risk patients.

4. Feasibility: H-4; M-11; L-1; 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 questioned the feasibility of collecting the data when only 1,016 SNFs currently use the CARE tool and there are about 15,326 SNFs across the country. The developer stated that the incorporation of the CARE tool into the MDS will make the data more feasible to collect.

3. Use and Usability: H-5; M-9; L-2; I-0

(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 be related to 2321: Functional Change: Change in Mobility Score and 2632: Long-Term Care Hospital (LTCH) Functional Outcome Measure: Change in Mobility Among Patients Requiring Ventilator Support. These measures have the same focus area (mobility) but are considered for different types of target populations. The Committee agreed that there was a need for all of the aforementioned measures; they made no recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-0**

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

Comments received:

- This measure received two comments in support for the measure, one of which mentioned the need for

2612 CARE: Improvement in Mobility
monitoring to ensure there would be no unintended consequences of the measure. An additional comment did not support the measure and raised concerns regarding unintended consequences around patient profiling. Committee response: • The Committee reviewed and discussed the comments on the post-comment committee call. Developer response: • AHCA thanks America's Health Insurance Plans for their comment. Any effective patient outcome quality measure has the potential to be utilized for patient profiling and this risk is minimized through the use of risk adjustors and exclusions. • AHCA thanks the Children's Hospital Association for their comment. We provided basic information on the measure inclusions within the measure specifications; however, we wanted to make sure that all of the detail information regarding the exclusions could be accessible for those interesting in replicating this measure. Therefore, we chose to place this more detailed information in the appendix.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2613 CARE: Improvement in Self Care
Submission |
Description: The measure calculates a skilled nursing facility's (SNFs) average change in self care for patients admitted from a hospital who are receiving therapy. The measure calculates the average change in self care score between admission and discharge for all residents admitted to a SNF from a hospital or another post-acute care setting for therapy (i.e., PT or OT) regardless of payor status. This is a risk adjusted outcome measure, based on the self care subscale of the Continuity Assessment and Record Evaluation (CARE) Tool and information from the admission MDS 3.0 assessment. The measure is calculated on a rolling 12 month, average updated quarterly. **Numerator Statement**: This outcome measure assesses the change in self-care. The numerator is the risk adjusted sum of the change in the CARE Tool self care subscale items between admission and discharge for each individual admitted from a hospital or another post-acute care setting regardless of payor status and are receiving therapy (PT or OT) for any reason in a skilled nursing center. **Denominator Statement**: The denominator includes all residents admitted to a SNF from a hospital or another post-acute care setting who receive either PT or OT therapy for any reason during their stay regardless of payor status, have a completed self care subscale of the CARE Tool at admission and discharge and do not meet any of the exclusion criteria and do not have missing data. The self care items used from the CARE tool are listed below and rated on a 1-6 scale (see Appendix for CARE Tool). The items included in the CARE Tool self care subscale include: • A1. Eating • A3. Oral Hygiene • A4. Toilet Hygiene • A5. Upper Body Dressing • A6. Lower Body Dressing • C1. Wash Upper Body • C2. Shower / Bathe • C6. Putting on / taking off footwear **Exclusions**: Individual patients are excluded for two broad reasons:

2613 CARE: Improvement in Self Care

1. if they have conditions where improvement in self-care is very unlikely,
OR
2. have missing data necessary to calculate the measure
Additionally, facilities with denominator size of fewer than 30 patients during a 12 month period are excluded from reporting of their data.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility

Type of Measure: Outcome

Data Source: Electronic Clinical Data, Other

Measure Steward: American Health Care Association

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-5**; **M-9**; **L-1**; **I-1;** 1c. High Priority: **H-10**; **M-5**; **L-1**; **I-0**

<u>Rationale</u>:

- The Committee agreed that the rationale supports a health outcome (change of self-care) in relation to the intervention of therapeutic services which are provided within the SNF; however, the Committee was unclear as to what a meaningful change in function would actually be for these patients, and how the measure as proposed relates to the quality of care provided and patient outcomes such as returning home (i.e., basic mobility skills, ADLs and IADLs activities).
- As with measure 2612, the Committee expressed major concerns about the measure focus on "improvement" when the Jimmo v. Sebelius settlement ruled that improvement in function cannot be required as a condition of coverage in SNFs (as well as home health and outpatient services). Therefore, without appropriate risk adjustment, a SNF may be more likely to deny access to patients who require SNF services to maintain or prevent further deterioration of function.
- The Committee noted that there were significant variation and room for improvement in terms of performance gap.
- The Committee further noted that specific data on disparities was not included in this measure as specified by current NQF requirements. The developer stated that to capture ethnicity, they would need to stratify, not risk adjust, for ethnicity, which would result in excluding over three-quarters of the SNFs in the country from this measure.
- The Committee was interested in in the inclusion of questions related to cognitive function. The developer stated that they risk-adjust for cognitive status using the MDS, however, if the CARE tool was inserted in the MDS, OASIS, and IRF-PAI, they would be better able to collect patients overall cognitive status.

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-7**; **M-6**; **L-1**; **I-2** 2b. Validity: **H-5**; **M-10**; **L-0**; **I-1**

<u>Rationale</u>:

- The Committee determined that the measure specifications were precise, noting that the specifications were consistent with the evidence presented.

2613 CARE: Improvement in Self Care

- Empiric reliability testing was performed at the patient level using data from the CARE tool. The Committee noted that reliability at the facility level was not provided and would be beneficial for the Committee to see the inter-classical correlation coefficients, a thumb print that suggests that there is a reproducible score within facilities that can then be used to compare between facilities.
- Empiric validity testing was conducted comparing the mobility measure set to other SNF quality measures including the 5 star rating, the Nursing Home Compare, and some specific measures like pressure ulcers and rehospitalization. There was no correlation between the mobility measure and the 5 star rating, but there were variable correlations between the mobility measure and Nursing Home Compare. The Committee was specifically interested in the 30 day SNF risk adjusted rehospitalization rates, where the developer's hypothesis (i.e., improved mobility leads to lower rehospitalization), did not come to fruition. Instead, validity testing showed that higher rehospitalization rates were positively correlated with a higher mobility score.
- The Committee supported the developer's decision to exclude specific patients (e.g., ventilator patients, persistent coma, quadriplegic, hospice, and children) from the measure, given certain concerns regarding failure to improve, as well as other unintended consequences associated with treating certain high-risk patients.

4. Feasibility: H-4; M-11; L-1; 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 questioned the feasibility of collecting the data when only 1,016 SNFs currently use the CARE tool and there are about 15,326 SNFs across the country. The developer stated that the incorporation of the CARE tool into the MDS will make the data more feasible to collect.

3. Use and Usability: H-5; M-9; L-2; I-0

(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 be related to 2286: Functional Change: Change in Self Care Score. These measures have the same focus area (self-care) but are considered for different types of target populations. The Committee agreed that there was a need for both measures; they made no recommendations for harmonization.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-0**

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

Comments received:

- This measure received two comments in support and two critical comments that raised concerns, one which explicitly did not support the measure and one of which did not explicitly state whether or not the commenter supported the measure. The focus of the concerns centers around the risk of unintended consequences around patient profiling. In addition, one of the critical comments raised additional concerns with the measure:

2613 CARE: Improvement in Self Care

"We continue to be concerned that the Improvement in Self-Care measures appears to consider self-care related movement alone and does not consider performance and cognitive elements of self-care such as sequencing, problem solving, temporal appropriateness (e.g., whether to dress for day or bed), memory, and activity planning. Further, it is notable that the Improvement in Self-Care measure does not consider or measure performance of activities of daily living, including the broader instrumental activities of daily living (IADLs) which significantly impact a patient's ability to function and live independently in the community."

NQF response:

- NQF is not able to monitor for unintended consequences directly, but we do encourage people to submit information via the Quality Positioning System to us should this problem arise. In addition, this comment has been forwarded to the developer.

Developer response:

- AHCA thanks the Children's Hospital Association for their comment. We provided basic information on the measure inclusions within the measure specifications; however, we wanted to make sure that all of the detail information regarding the exclusions could be accessible for those interesting in replicating this measure. Therefore, we chose to place this more detailed information in the appendix.
- AHCA thanks America's Health Insurance Plans for their comment. Any effective patient outcome quality measure has the potential to be utilized for patient profiling and this risk is minimized through the use of risk adjusters and exclusions.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2624 Functional Outcome Assessment

Submission |

Description: NOTE: Specification information in this section is from the 2014 Physician Quality Reporting System Manual. Note that Testing Information is based on the specification in the 2012 Physician Quality Reporting System Manual. Both 2012 and 2014 Specifications are included in the attached "NQF Endorsement Measurement Submission Summary Materials"

Percentage of visits for patients aged 18 years and older with documentation of a current functional outcome assessment using a standardized functional outcome assessment tool on the date of the encounter AND documentation of a care plan based on identified functional outcome deficiencies on the date of the identified deficiencies.

Numerator Statement: Patients with a documented current functional outcome assessment using a standardized tool AND a documented care plan based on the identified functional outcome deficiencies.

Denominator Statement: All visits for patients aged 18 years and older

Exclusions: Not Eligible – A patient is not eligible if one or more of the following reasons(s) is documented:

• Patient refuses to participate

• Patient unable to complete questionnaire

• Patient is in an urgent or emergent medical situation where time is of the essence and to delay treatment would jeopardize the patient's health status

Adjustment/Stratification:

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Ambulatory Care : Outpatient Rehabilitation

2624 Functional Outcome Assessment

Type of Measure: Process
Data Source: Administrative claims, Paper Medical Records
Measure Steward: Centers for Medicare and Medicaid Services

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 Priority)
1a. Evidence: **H-2**; **M-9**; **L-3**; **I-3; IE-2**; 1b. Performance Gap: **H-4**; **M-12**; **L-3**; **I-0;** 1c. High Priority: **H-4**; **M-12**; **L-3**; **I-0**
<u>Rationale</u>:

- The committee questioned the evidence base showing that the documentation of a standardized functional assessment care planning improves patient outcomes. The developer agreed that while there is definitely an established link between the care itself and the outcome measure, the evidence linking a recording of the use of the tool and improved outcomes is less strong.
- The median performance rate on this measure is 100%, but the developer states that this is based on a very small, self-selected population and that that data needs to be taken with a grain of salt. The developer's data indicates that the 2012 average performance rate on this measure was 80.9 percent, and that there were differences in performance rates among various demographic groups. In 2013, more providers were reporting on this measure but they do not have updated averages and medians yet. They noted that the low number of participating providers could also indicate a gap in care. Developers noted that this process measure is intended to be an intermediate step to a future outcome measure.
- Committee members agreed that assessing function and developing a plan of care are basic practices for PT, OT, and chiropractic providers.

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-4**; **I-5 (consensus not reached)** 2b. Validity: **H-0**; **M-8**; **L-6**; **I-5 (consensus not reached)**
UPDATED VOTES FOR 2a. Reliability: H-0; **M-17**; **L-1**; **I-1 2b. Validity: H-1**; **M-15**; **L-2**; **I-1**
<u>Rationale</u>:

- The developers confirmed that both parts of the measure must be passed in order to meet the measure.
- Committee members wanted more information on operationalizing the measure; specifically, how to ensure a documented care plan would be addressing the identified functional outcome deficiencies and how that would be coded. The developer explained that they do not need to be linked but that both a functional status assessment and a care plan need to show up in the record.
- Committee members were also concerned that using claims data to fulfill this measure does not link the two pieces of the measure when it comes to actual provision of care. Committee members were concerned this measure is very "game-able" as the documentation of a care plan would fulfill the measure, but would not ensure that the patient received the right care. The developers noted that linking the care plan and the collection of outcomes data would naturally be linked for providers.

4. Feasibility: H-3; M-11; L-5; 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>:

- Committee members noted that this is based on claims data; however, only 3.6% of eligible providers

2624 Functional Outcome Assessment
reported it, which could indicate feasibility issues. Data is abstracted from administrative claims and paper medical records, which can reduce feasibility.
3. Use and Usability: H-4; M-9; L-6; I-0 *(Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting/Accountability and 3b. Quality Improvement)* Rationale: • The measure is currently in use in PQRS.
5. Related and Competing Measures • The Committee considered this measure to be related to the set of FOTO measures on functional status change (0422, 0423, 0424, 0425, 0426, 0427, 0428), as they address functional status for patients age 18 years and older. However, they differ in type as 2624 is a process measure while FOTO measures assess patient-reported outcomes. The Committee agreed that these measures were related but did not make recommendations for harmonization. •
UPDATED VOTES FOR Standing Committee Recommendation for Endorsement: Y-10; N-9(consensus not reached); Y-17; N-2
6. Public and Member Comment : March 2, 2015- March 31, 2015 Comments received: • There was only one comment received on this measure which supported its endorsement. NQF response: • Thank you for your comment.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2631 Percent of Long-Term Care Hospital (LTCH) Patients With an Admission and Discharge Functional Assessment and a Care Plan That Addresses Function
Submission |
Description: This quality measure reports the percentage of all Long-Term Care Hospital (LTCH) patients with an admission and discharge functional assessment and a care plan that addresses function. **Numerator Statement**: The numerator for this quality measure is the number of Long-Term Care Hospital (LTCH) patients with complete functional assessment data and at least one self-care or mobility goal. For patients with a complete stay, all three of the following are required for the patient to be counted in the numerator: (1) a valid numeric score indicating the patient's status or response, or a valid code indicating the activity was not attempted or could not be assessed, for each of the functional assessment items on the admission assessment; (2) a valid numeric score, which is a discharge goal indicating the patient's expected level of independence, for at least one self-care or mobility item on the admission assessment; and (3) a valid numeric score indicating the patient's status or response, or a valid code indicating the activity was not attempted or could not be assessed, for each of the functional assessment items on the discharge assessment. For patients who have an incomplete stay, discharge data are not required. The following are required for the patients who have an incomplete stay to be counted in the numerator: (1) a valid numeric score indicating the patient's status or response, or a valid code indicating the activity was not attempted or could not be assessed, for each of the functional assessment items on the admission assessment; and (2) a valid numeric score, which is a

2631 Percent of Long-Term Care Hospital (LTCH) Patients With an Admission and Discharge Functional Assessment and a Care Plan That Addresses Function

discharge goal indicating the patient's expected level of independence, for at least one self-care or mobility item on the admission assessment.

Patients who have incomplete stays are defined as those patients (1) with incomplete stays due to a medical emergency, (2) who leave the LTCH against medical advice, or (3) who die while in the LTCH. Discharge functional status data are not required for these patients because these data may be difficult to collect at the time of the medical emergency, if the patient dies or if the patient leaves against medical advice.

Denominator Statement: The denominator is the number of LTCH patients discharged during the targeted 12 month (i.e., 4 quarters) time period.

Exclusions: There are no denominator exclusions for this measure.

Adjustment/Stratification:

Level of Analysis: Facility

Setting of Care: Post Acute/Long Term Care Facility : Long Term Acute Care Hospital

Type of Measure: Process

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: **H-1**; **M-5**; **L-9**; **I-9; IE-3**; 1b. Performance Gap: **H-1**; **M-2**; **L-3**; **I-12;** 1c. High Priority: **H-0**; **M-0**; **L-0**; **I-0**

UPDATED VOTES FOR 1a. Evidence: H-4; M-10; L-1; I-0; IE-2; (pass) 1b. Performance Gap: H-2; M-8; L-3; I-4; (consensus not reached) 1c. High Priority: H-10; M-6; L-1; I-0 (pass)

UPDATED VOTES FOR 1b. Performance Gap: H-4; M-11; L-2; I-1

<u>Rationale</u>:

- This measure was not recommended initially in the January 21-22 Committee in-person meeting because it did not pass the importance criteria. However, the Committee conducted a subsequent review of this measure on the January 28 post-meeting call per the developer's request. This time the measure passed the importance criteria in the gray zone based on the additional information that was presented by the developers.
- The developer noted that this measure has two components, including: 1) the collection of standardized functional assessment data in the areas of self-care, mobility, cognition, and bladder management, and 2) the reporting, on admission, of a discharge goal (i.e., score) for one or more self-care or mobility items.
- The Committee questioned whether the two components of documenting a functional status assessment on admission and a goal for function are linked together. The developer responded that the goal has to be tied to one of the self-care or mobility items. So if the person has a functional limitation in eating, rolling left or right, getting on and off the toilet, the clinicians have to report a goal for at least one of those items using the functional scale.
- The Committee expressed concern that the only data presented to support the performance gap was qualitative data collected from site visits to 28 facilities and there were no quantitative data and data for a care plan gap. The developer stated that based on their understanding, qualitative data is sometimes adequate for a measure when it is first being proposed especially process measures that are directly tied to expert opinion in terms of validity and clinical practice guidelines.

2631 Percent of Long-Term Care Hospital (LTCH) Patients With an Admission and Discharge Functional Assessment and a Care Plan That Addresses Function

- The Committee also had concerns that this might be a hard measure to get a good grade on because there are three components to the numerator which the long-term care facilities have to comply with. The developer explained that all the items will be nested within the LTCH care data set and collected through a standardized assessment tool, which long-term hospitals are required to use.
- Additional questions were raised regarding the assessment in setting a goal for the purpose of data collection versus holding the facility accountable for that goal. The developer explained that CMS is attempting to collect data to examine a change in independence on self-care and mobility and see if these items line up to a goal of care and then standardize data assessment across settings to follow persons as they traverse across care settings.

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-0**; **L-0**; **I-0** 2b. Validity: **H-0**; **M-0**; **L-0**; **I-0**
UPDATED VOTES FOR 2a. Reliability: H-0; **M-7**; **L-5**; **I-5** (**consensus not reached) 2b. Validity: H-0; M-7; L-4; I-6 (consensus not reached)**
UPDATED VOTES FOR 2a. Reliability: H-4; M-12; L-2; I-0 2b. Validity: H-2; M-14; L-2; I-0
<u>Rationale</u>:

- The Committee noted that there is a good evidence for reliability and validity of the care component and the functional status component, but there is no data regarding the care plan piece of the measure. The measure also lacks the inter-rater reliability data on the degree to which an appropriate goal is set. The developer responded that the "appropriate" in this argument may not essentially fit within this measure. This measure is just looking at the items for self-care and mobility and whether one of those items was documented on the goal of care at discharge.
- One Committee member raised concerns about the face validity of the measure if documentation of functional status and a related goal is called a care plan. Another Committee member agreed that a goal is not equal to a care plan; however, she supported the idea of a measure that links current functional status and the goal for improvement and suggested the developer tweak the semantics for this measure. CMS will consider revising the measure title to address the Committee's concerns.
- One Committee member pointed out that there is no evidence of intraclass correlation coefficients that would suggest the signal to noise ratio which helps distinguish within facility variability from between facility variation and asked the developers whether they have the data to analyze that. The developer explained that they don't have data to analyze facilities over time. As part of the post-acute payment reform demonstration, they had 28 LTCHs volunteered to use the standardized dataset to collect and enter data into an electronic system whereby provided the reliability and validity data.

4. Feasibility: H-0; M-0; L-0; I-0
UPDATED VOTES FOR Feasibility: H-4; M-12; L-1; 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>:

- All data elements are in defined fields in electronic clinical data and the functional assessment items included in this quality measure will be included in a future version of the LTCH CARE Data Set (Version

2631 Percent of Long-Term Care Hospital (LTCH) Patients With an Admission and Discharge Functional Assessment and a Care Plan That Addresses Function

3.00). The LTCH CARE Data Set has been the assessment data set used in LTCHs since 2012, when the LTCH Quality Reporting Program was implemented, as required by the Patient Protection and Affordable Care Act.

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

UPDATED VOTES FOR Use and Usability: H-2; M-12; L-3; I-0

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

Rationale:

- Data collection for this quality measure begins on April 1, 2016 as part of the Long-Term Care Hospital Quality Reporting Program. Proposed plans for the public reporting of this quality measure will be included in future rulemaking published in the Federal Register.
- A Committee member raised a question regarding the possibility of non-response rate of facilities in terms of reporting this data. CMS explained that LTCHs that do not collect and submit data for this measure by the submission deadline may be subject to a two percentage point reduction in the annual payment update for fiscal year 2018 and subsequent years.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: Y-9; N-8 (consensus not reached)Y-15; N-3

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

Comments received:

- Two comments were received on this measure of which one supported the endorsement of this measure. The second commenter noted that this measure is an important topic within the PAC industry and has been subject to contentious discussions across NQF committees and raised concern about the NQF processes for re-consideration and re-voting. The commenter further noted that the MAP Committees have "Conditionally Supported" this measure for use within all PAC venues and recommended that the Committee take all PAC settings into consideration when reviewing this measure to identify whether it meets all of the criteria previously reviewed not just for LTCHs, but also for SNFs, IRFs, and Home Health agencies.

NQF response:

- We appreciate your input, but would note that this measure was re-discussed during the follow up call after the in-person meeting. During the meeting, the Committee requested additional information regarding the measure. The developers had already submitted this information; however, due to timing of receipt being just prior to the in-person meeting; the Committee did not have time to review it. Due to the fact the information was already available, NQF agreed to have the Committee re-discuss the measure during the post-meeting call rather than waiting until after the public comment period.

Committee response:

- This comment was addressed on the post-comment call. Consensus has not been reached on some of the required criteria, and additional information was requested. While the comments on expanding the settings for the measure's use are appreciated, the Committee is charged with evaluating measures based on the information submitted and for the level of analysis and care setting as submitted by the developer. This measure was recommended by the Committee after reviewing the additional information and the comments.

Developer response:

- Thank you for your comment. 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

2631 Percent of Long-Term Care Hospital (LTCH) Patients With an Admission and Discharge Functional Assessment and a Care Plan That Addresses Function
cognitive function. Following the enactment of the IMPACT Act, a technical expert panel (TEP) was convened by the Centers for Medicare and Medicare Services' measure development contractor and provided input on implementing an application of this measure across four post-acute care settings, including IRFs, LTCHs, SNFs and HHAs. The TEP supported the implementation of this measure as specified across PAC providers and also supported our efforts to standardize this measure for cross-setting use. The Measures Application Partnership (MAP) met on February 9, 2015 and conditionally supported the specification of an application of Percent of LTCH Patients With an Admission and Discharge Functional Assessment and a Care Plan That Addresses Function (NQF #2631; under review) for use as a cross-setting measure. MAP conditionally supported this measure pending NQF-endorsement and resolution of the use of two different functional status scales for quality reporting and payment purposes. MAP reiterated its support for adding measures addressing function, noting the group's special interest in this PAC/LTC core concept. More information about the MAPs recommendations for this measure is available at: http://www.qualityforum.org/Setting_Priorities/Partnership/MAP_Final_Reports.aspx.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2632 Long-Term Care Hospital (LTCH) Functional Outcome Measure: Change in Mobility Among Patients Requiring Ventilator Support
Submission |
Description: This measure estimates the risk-adjusted change in mobility score between admission and discharge among LTCH patients requiring ventilator support at admission. **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 LTCH patients requiring ventilator support at admission. The change in mobility score is calculated as the difference between the discharge mobility score and the admission mobility score. **Denominator Statement**: The target population (denominator) for this quality measure is the number of LTCH patients requiring ventilator support at the time of admission to the LTCH. **Exclusions**: 1) Patients with incomplete stays: Rationale: It can be challenging to gather accurate discharge functional assessment data for patients who experience incomplete stays. Patients with incomplete stays include patients who are unexpectedly discharged to an acute-care setting (Inpatient Prospective Payment System or Inpatient Psychiatric Hospital) because of a medical emergency or psychiatric condition; patients transferred to another LTCH facility; patients who leave the LTCH against medical advice; patients who die; and patients with a length of stay less than 3 days. 2) Patients discharged to hospice: Rationale: Patients discharged to hospice are excluded because functional improvement may not be a goal for these patients. 3) Patients with progressive neurological conditions, including amyotrophic lateral sclerosis, multiple sclerosis, Parkinson's disease, and Huntington's chorea: Rationale: These patients are excluded because they may have functional decline or less predictable function trajectories. 4) Patients in coma, persistent vegetative state, complete tetraplegia, and locked-in syndrome: Rationale: The patients are excluded because they may have limited or less predictable mobility recovery.

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

5) Patients younger than age 21:
Rationale: There is only limited evidence published about functional outcomes for individuals younger than 21.
6) Patients who are coded as independent on all the CARE mobility items at admission:
Rationale: These patients are excluded because no improvement in mobility skills can be measured with the mobility items used in this quality measure.
Adjustment/Stratification:
Level of Analysis: Facility
Setting of Care: Post Acute/Long Term Care Facility : Long Term Acute Care Hospital
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: The measure meets the Importance criteria
(1a. Evidence, 1b. Performance Gap, 1c. High Impact)
1a. Evidence: **Y-13**; **N-0**; 1b. Performance Gap: **H-1**; **M-8**; **L-0**; **I-4;** 1c. High Priority: **H-3**; **M-10**; **L-0**; **I-0**
Rationale:

- The developer stated that functional improvement is particularly relevant for patients who require ventilator support, since these patients traditionally have limited or no mobility because of cardiovascular and pulmonary instability, delirium, sedation, lack of rehabilitation therapy staff, and lack of physician referral. The Committee appreciated the background and especially the note that the measure is required by law, and is an example of where CMS is moving toward standardization and alignment of measures.
- The Committee noted the small study sample of 103 patients with respiratory failure needing ventilator support; however the developer corrected that the sample size was actually 455 patients. The Committee understands the complexity of getting a big enough sample for this type of patient population, however, expects that over time the developer will garner more adequate data.
- The Committee also noted the lack of data on performance gap. The developer explained that there is not a lot of literature about long-term care hospitals patients, in particular ventilator patients.

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-1**; **M-7**; **L-3**; **I-2** 2b. Validity: **H-0**; **M-9**; **L-2**; **I-2**
Rationale:

- As with previous metrics, data was provided at the item or scale level which is acceptable under NQF critieria; the Committee noted lack of facility level testing data.
- The Committee questioned the developer's decision to utilize a complex calculation for the risk adjusted change score when there are other ways of risk adjusting that does not involve a predictive score. The Committee expressed concerns about reliability testing using the predictive score as opposed to the actual score.
- The Committee expressed concerns about the exclusion of all progressive neurologic conditions, especially MS and Parkinson's disease, where the patients' ability to function on or off a ventilator often fluctuates. While the Committee understands the impetus for excluding that population, they questioned

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

what proportion of the 300,000 ventilated patients annually fall under those categories and which categories. Additionally, there might be some progressive neurologic conditions that would require exclusion (e.g., ALS) however their might be others that should be included (e.g., MS, Parkinson's). The Committee encouraged further exploration of the data to unearth some of these progressive neurologic conditions that should be included. The developer indicated inclusion of data and impact of exclusions in their submission and also indicated their technical expert panel considered exclusion at length during the development of the measure.

- The Committee questioned the developer's decision to omit patients who are on ventilators at home, nursing homes, and inpatient rehab facilities, when the goal is to standardize measures across all care settings. The developer indicated that the vast majority of PAC patients on ventilators are in LTCH's which are the care level included in the law that is the impetus for this measure. It was noted there would be a sample size issue if looking at patients on ventilators in other settings. The Committee agreed this would be an issue but still may be worth evaluating: where the patient is best served, being able to transport patients from one level of care to another and sharing information. The Committee understands the developers reasoning the exclusion however voiced that these populations should be included over time once the measure gets implemented.

4. Feasibility: H-3; M-9; L-1; 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 raised no concerns with the measure's feasibility.

3. Use and Usability: H-3; M-7; L-1; I-2

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

Rationale:

- The Committee raised no concerns with the measure's use or usability.

5. Related and Competing Measures

The Committee considered this measure to be related to 2612: CARE: Improvement in Mobility and 2321: Functional Change: Change in Mobility Score. 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.

Standing Committee Recommendation for Endorsement: **Y-12**; **N-1**

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

Comments received:

- One commenter stated that there does not appear to be a specific age exclusion for this measure and inquired whether the measure has been tested in patients under the age of 18.

Developer response:

- Thank you for your comment. Our testing data included patients in long-term care hospitals who were 20 to 99 years old. It did not include patients who were 18 or younger. However, we would like to note that this is a process measure focused on whether a functional assessment is completed and whether a functional goal is reported. It is not an outcome measure, and does not include comparing patient scores. The objective of this measure is to promote standardized functional assessment of basic daily activities for

2632 Long-Term Care Hospital (LTCH) Functional Outcome Measure: Change in Mobility Among Patients Requiring Ventilator Support
all patients. Therefore, we believe it applies to all patients, regardless of age.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals

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

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

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

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

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, the Committee did not achieve consensus on whether one measure is superior. Therefore, these measures will go forward to NQF Member vote as recommended and will be brought forward to the CSAC on June 9th for resolution of the best in class decision.
- 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; 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

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

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. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

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: The measure meets the Importance criteria

(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**

Rationale:

- 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

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

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 potentially compete with 2321: Functional Change: Change in Mobility Score (UDSMR) 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 is superior. Therefore, these measures will both continue on to NQF member vote as recommended and be brought forward to the CSAC on June 9th for resolution of the best-in-class decision.
- 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.

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

-

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. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

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

Submission |

Description: This measure estimates the percentage of IRF patients who meet or exceed an expected discharge self-care score.

Numerator Statement: The numerator is the number of patients in an IRF with a discharge score that is equal to or higher than the calculated expected discharge score.

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

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

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 discharged to a hospice; patients discharged to another IRF; 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 with the following medical conditions: coma; persistent vegetative state; complete tetraplegia; 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.

3) Patients younger than age 21.

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

4) Patients discharged to Hospice.

Rationale: Patient goals may change during the IRF stay.

5) 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 prepared 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 for the measure's evidence. 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 outcome measurement, whether it is self-care or mobility.
- The Committee asked for the reasoning behind the proposal of four measures using essentially the same data. The developer explained 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 in the same

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

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; it could be used internally for the facility for quality improvement and externally 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 the 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-11**; **L-2**; **I-0 2b. Validity: H-3**; **M-14**; **L-0**; **I-1**
<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 noted 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 gameable. 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 was submitted late and may have come in after the Committee reviewed each measure. The 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 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 a standing committee, that data will come back 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)
<u>Rationale</u>:

- 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

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

(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 be related to 2633: IRF Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients (CMS) and 2286: Functional Change: Change in Self-Care Score (USDMR), however there were no recommendations for harmonization.

Standing Committee Recommendation for Endorsement: Y-10; N-5; Y-17; N-1

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

2635 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Self-Care Score for Medical Rehabilitation Patients
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. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2636 Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Discharge Mobility Score for Medical Rehabilitation Patients
Submission |
Description: This measure estimates the percentage IRF patients who meet or exceed an expected discharge mobility score. **Numerator Statement**: The numerator is the number of patients in an IRF with a discharge mobility score that is equal to or higher than a calculated expected discharge mobility score. **Denominator Statement**: IRF patients included in this measure are at least 21 years of age, Medicare beneficiaries, and have complete stays. **Exclusions**: This quality measure has 4 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 with the following medical conditions on admission: 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 items. 3) Patients younger than age 21. Rationale: There is only limited evidence published about functional outcomes for individuals younger than 21. 4) Patients discharged to hospice. Rationale: Patient goals may change during the IRF stay. 5) Patients who are not Medicare beneficiaries. **Adjustment/Stratification**: **Level of Analysis:** Facility

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

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: The measure meets the Importance criteria
(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**
Rationale:

- 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 reporting 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 quality data however it is not tied to a pay-for-performance program.

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

- The developers utilized different types of reliability including inter-rater reliability and the use of video to assess clinician assessments. 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' change in function.
- The Committee expressed concerns that reliability and validity data was at the scale level and the not 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 the observed or expected calculation comes from the other 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)
Rationale:

- 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)
Rationale:

- The developer noted that IRF measures are limited to Medicare only and that the Long-Term Care

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

Hospital Quality Reporting Program was established as a Medicare program.

5. Related and Competing Measures

- The Committee considered this measure to be related to 2321: Functional Change: Change in Mobility Score and 2634: Inpatient Rehabilitation Facility (IRF) Functional Outcome Measure: Change in Mobility Score for Medical Rehabilitation Patients. 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. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2643 Average change in functional status following lumbar spine fusion surgery

Submission |

Description: For patients age 18 and older undergoing lumbar spine fusion surgery, the average change from pre-operative functional status to one year (nine to fifteen months) post-operative functional status using the Oswestry Disability Index (ODI version 2.1a) patient reported outcome tool.

Numerator Statement: There is not a traditional numerator for this measure; the measure is calculating the

2643 Average change in functional status following lumbar spine fusion surgery

average change in functional status score from pre-operative to post-operative functional status score. The measure is NOT aiming for a numerator target value for a post-operative ODI score.

For example:

The average change in low back function was an increase in 17.2 points one year post-operatively on a 100 point scale.

Denominator Statement: Adult patients age and older (no upper age limit) who undergo a lumbar spine fusion procedure during a calendar year performance period (e.g. dates of procedure occurring between 1/1/2013 and 12/31/2013) AND have a completed pre-operative and post-operati

Exclusions: Exclusions are for patients with spine related cancer, fracture and infection and idiopathic or congenital scoliosis.

Adjustment/Stratification:

Level of Analysis: Clinician : Group/Practice

Setting of Care: Ambulatory Care : Clinician Office/Clinic

Type of Measure: PRO

Data Source: Electronic Clinical Data : Electronic Health Record, Paper Medical Records, Patient Reported Data/Survey

Measure Steward: MN Community Measurement

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 Priority)

1a. Evidence: **Y-18**; **N-1**; 1b. Performance Gap: **H-6**; **M-8**; **L-0**; **I-5** 1c. High Priority: **H-13**; **M-6**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer introduced this new measure as a patient-reported outcome measure, which evaluates the change between a patient's preoperative functional status and their post-operational functional status at one year.
- The Committee applauded the developers for tackling this controversial and important area in utilization of surgical procedures, pointing to the developer's statement that there is a 15-fold increase in the number of complex fusion procedures performed for Medicare beneficiaries, which is a highly variable procedure. However, the Committee stated that this measure could imply that there is a gap in quality of care, but not a gap in variability in performance, based on the pilot data.
- The developer explained that this measure has gone through one phase of pilot testing involving four practices and is in the statewide quality reporting and measurement system for Minnesota, which is required of practitioners. The developer noted that they are expecting full implementation data to be available in May, 2015.
- The Committee members raised additional concerns that the Oswestry tool (pain questionnaire) may not be the best tool to use, because it is primarily aimed at pain and therefore would not capture other neurological dysfunctions or potential side effects of the surgery itself. They recommended that the measure be improved by adding other questions or tools that might speak to neurological symptoms that could present without pain.

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-6; L-4; I-9 2b. Validity: H-X; M-X; L-X; I-X

2643 Average change in functional status following lumbar spine fusion surgery

UPDATED VOTES FOR 2a. Reliability: H-3; **M-15**; **L-0**; **I-0 2b. Validity: H-1**; **M-17**; **L-0**; **I-0**

Rationale:

- The Committee members commented that the specifications look very clear but the risk adjustment specifications have not been modeled yet. Further, the Committee noted that there is no score-level reliability testing data presented as well as data to demonstrate the intraclass correlations at the practice-level. The developer confirmed that similar to *Measure 2653 Average Change in Functional Status Following Total Knee Replacement Surgery,* they will submit testing based on the Committee's recommendations.
- This measure did not pass reliability so the Committee stopped voting at this juncture and requested the aforementioned testing information from the developers to re-consider the measure after the public comment. One Committee member offered an additional suggestion for the developers to add questions such as whether or not non-invasive treatments were tried (e.g., physical therapy or pain consults, steroid injections) to get a sense for onset of symptoms, other treatments that were tried, and clinical indications for the procedure.

3. Feasibility: H-3; M-14; L-1; I-0

(3a. Data generated during care; 3b. Electronic sources; and 3c. Data collection can be implemented (eMeasure feasibility assessment of data elements and logic)

Rationale:

- This measure data source is Electronic Clinical Data, Electronic Health Record, Paper Medical Records, and Patient Reported Data/Survey.
- Also, all data elements are in defined fields in electronic health records (EHRs)

4. Use and Usability: H-3; M-14; L-0; I-1

(4a. Accountability/transparency; and 4b. Improvement – progress demonstrated; and 4c. Benefits outweigh evidence of unintended negative consequences)

Rationale:

- This measure is not currently in use but planned for use in public reporting, payment program, and regulatory and accreditation programs.
- The developer also noted that this measure is planned for inclusion in the MN Department of Health (MDH) Statewide Quality Reporting and Measurement System. Mandatory data collection and reporting under 2008 MN Health Reform Legislation. MNCM was a subcontractor to MDH for measure development exploring the concept of low back pain. Statewide implementation is planned for submission in April/May 2015 for dates of procedure 1/1/2013 to 12/31/2013 with follow-up assessment period through March 31, 2015.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-X**; **N-X, Y-18**; **N-0**

6. Public and Member Comment

Comments received:

- Commenters believed this measure should be considered for endorsement once the reliability testing data is submitted by Minnesota Community Measurement because the measure focuses on an important patient-centered outcome and addressees an important gap area for quality improvement. We believe an explicit patient-centered focus on surgical outcomes is necessary and this measure begins to address this important quality issue.

Committee response:

2643 Average change in functional status following lumbar spine fusion surgery

- 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 support! We agree that these types of measures focused on patient reported outcomes and change over time, which represent newer cutting-edge measures, are more difficult to evaluate as compared to traditional measures that are expressed as a binary Yes/No. We have provided additional testing in response to the Standing committee's concerns and look forward to continued conversation and working with NQF staff to determine the best statistical methods and tests for determining the reliability and validity performance scores. A new published study supports the use of the Oswestry Disability Index as a PROM tool appropriate for outcome measurement. "A proposed set of metrics for standardized outcome reporting in the management of low back pain." Clement, RC et al Acta Orthopaedica 2015; 86 (4)

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2653 Average change in functional status following total knee replacement surgery

Submission |

Description: For patients age 18 and older undergoing total knee replacement surgery, the average change from pre-operative functional status to one year (nine to fifteen months) post-operative functional status using the Oxford Knee Score (OKS) patient reported outcome tool.

Numerator Statement: There is not a traditional numerator for this measure; the measure is calculating the average change in functional status score from pre-operative to post-operative functional status score. The measure is NOT aiming for a numerator target value for a post-operative OKS score.

For example:

The average change in knee function was an increase of 15.9 points one year post-operatively on a 48 point scale.

Denominator Statement: Adult patients age and older (no upper age limit) who undergo a primary or revision total knee replacement procedure during a calendar year performance period (e.g. dates of procedure occurring between 1/1/2013 and 12/31/2013) AND have a completed pre-operative and post-operative OKS patient reported outcome assessments.

Exclusions: There are no denominator exclusions from the initial patient population for this measure.

Adjustment/Stratification:

Level of Analysis: Clinician : Group/Practice

Setting of Care: Ambulatory Care : Clinician Office/Clinic

Type of Measure: PRO

Data Source: Electronic Clinical Data : Electronic Health Record, Paper Medical Records, Patient Reported Data/Survey

Measure Steward: MN Community Measurement

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

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

2653 Average change in functional status following total knee replacement surgery

(1a. Evidence, 1b. Performance Gap, 1c. High Impact)
1a. Evidence: **Y-16**; **N-3**; 1b. Performance Gap: **H-3**; **M-12**; **L-3**; **I-1;** 1c. Impact: **H-8**; **M-9**; **L-2**; **I-0**
Rationale:

- The number of total knee replacements (TKR) is rising and will continue to rise over the next 10 years as the Baby Boom generation ages, especially because the standard of care for end-stage degenerative arthritis of the knee is knee arthroplasty. The Committee agreed that patients could use information on what level of functional status they can expect after a TKR. The Committee agreed that a one-year postoperative assessment was the right timeframe as much sooner would not accurately measure real outcomes.
- The Committee requested more information on the effect of the measured improvement and whether is an amount that actually impacts outcomes. While the developers had not included that information in the submission form, a Committee member provided data from a different study that the standard deviation is 8, and so the difference noted in the measure would be very significant (two standard deviations).
- The Committee was concerned that the measure did not collect data on postoperative interventions, such as rehabilitation, that could affect outcomes separately from the surgery and that could not be held attributable to the surgeon. However, one Committee member suggested that a surgeon should be seen as the leader of a team taking care of a knee and that this sort of measure would encourage more focus on long-term outcomes.
- The developer confirmed that patients were involved in the measure development work group.

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-0**; **M-9**; **L-3**; **I-7** (consensus not reached) 2b. Validity: **H-1**; **M-7**; **L-5; I-6 (consensus not reached)**
UPDATED VOTES FOR 2a. Reliability: H-2; **M-13**; **L-3**; **I-0 2b. Validity: H-2**; **M-14**; **L-1; I-1**
Rationale:

- Committee members questioned why the measure is not risk adjusted; the developer explained that they plan to test a number of variables when they have more data as potential adjustors.
- The Committee wanted to know how different the average patient population in Minnesota would be from the average patient across the US, and also raised concerns that the measure was originally tested on a very different patient population, potentially affecting the reliability and validity of the instrument used.
- The developer stated that this is a new type of measure that does not have a traditional numerator and denominator; it is a continuous measure. They stated that measurement science has not yet evolved to the point of determining appropriate methodology for testing reliability for this type of measure. The Committee suggested intraclass correlate testing as a possibility.
- The developer mentioned they had some difficulty with the PRO tool administration rates and that they were working with a phased approach to improve those rates.
- Committee members requested an estimation of reliability at the physician level and the developers agreed to follow up with that information.

4. Feasibility: H-1; M-15; L-2; I-1

2653 Average change in functional status following total knee replacement surgery

(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:

- In response to a request for more information, the developers explained that the measure requires a pre-operative OKS summary score, using a simple tool. Practices submit patient-level information to a portal that calculates the measures. They noted that orthopedic practices are new to measurement, and that their pilot groups said getting information into and out of their EMRs was easier than getting the patient-reported tools into their workflows. However, they are seeing gradual improvements.
- Post-operatively, the tool is filled out during an office visit or sent to the patient via mail or the patient portal.
- Committee members asked if this measure is susceptible to gaming; the developers said that there are no appropriateness criteria guidelines for knee replacement but that they collect the data on all patients, whether they completed one or both assessments.

3. Use and Usability: H-0; M-12; L-5; I-2

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

Rationale:

- The measure is not currently in use, so no usability data is available, but the developer plans to report it statewide in Minnesota in 2016.

5. Related and Competing Measures

- The Committee considered whether this measure potentially competes with 0422: Functional Status Change for Patients with Knee Impairments (FOTO). The Committee determined that the measures have different focus in terms of the target population, provider types, and clinical settings, as well as the clinical area. The developers indicated that the FOTO measure is broader and applicable to any kind of knee impariments, as opposed to measure 2653, which only focuses on patients with knee replacements. Therefore, the Committee agreed that the measures were related but not competing. The Committee did not make recommendations for harmonization.

Standing Committee Recommendation for Endorsement: Y-11; N-8 (consensus not reached); Y-15; N-3

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

Comments received:

- Commenters strongly urge the Committee to reconsider and recommend this measure. The measure is deemed by consumers and purchasers to be important for assessing providers of knee replacement surgery. This is a high frequency and high cost procedure, and currently there is no information that enables patients to choose providers that can achieve better outcomes as assessed by patients themselves. Therefore, this measure is a high priority for these users. Commenters also asked NQF to consider ways to improve upon the validity and reliability of this measure and other similar measures should be considered in the future.

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.

NQF response:

- NQF has reviewed your comment and appreciates your input. Your comment has been forwarded to the

2653 Average change in functional status following total knee replacement surgery
Standing Committee and Developer for consideration. NQF is not able to improve measures as our role is to endorse measures, not maintain them, but we do encourage improvements to measures over time and at the three-year maintenance cycle review. Developer response: • Thank you for your support! We agree that these types of measures focused on patient reported outcomes and change over time, which represent newer cutting-edge measures, are more difficult to evaluate as compared to traditional measures that are expressed as a binary Yes/No. We have provided additional testing in response to the Standing committee's concerns and look forward to continued conversation and working with NQF staff to determine the best statistical methods and tests for determining the reliability and validity performance scores. Thanks for your suggestion to determine modes that address survey burden. In addition to obtaining survey information from the patient during an in-person visit, we do allow mailed survey and when permitted by the tool developer/ copyright holder, electronic administration of the tool to the patient by patient portal. Additionally, although not yet submitted for endorsement, MN Community Measurement is also measuring the change in quality of life for this patient population, initially using the EQ5D and now transitioning to PROMIS Global Health-10. • We agree that these types of measures focused on patient reported outcomes and change over time, which represent newer cutting-edge measures, are more difficult to evaluate as compared to traditional measures that are expressed as a binary Yes/No. We have provided additional testing in response to the Standing committee's concerns and look forward to continued conversation and working with NQF staff to determine the best statistical methods and tests for determining the reliability and validity performance scores.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X; A-X
8. Board of Directors Vote: Y-X; N-X
9. Appeals