NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Click to go to the link. ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 3325

Measure Title: Long Term Services and Supports (LTSS) Shared Care Plan with Primary Care Practitioner

Measure Steward: Centers for Medicare & Medicaid Services, Centers for Medicaid & CHIP Services

Brief Description of Measure: This measure assesses the percentage of Medicaid Managed Long Term Services and Supports (MLTSS) Plan enrollees with a care plan for whom all or part of the care plan was transmitted to the primary care practitioner (PCP) within 30 days of the care plan's development or update.

Developer Rationale: This measure would address the lack of coordination between LTSS and medical care providers by ensuring a patient's care plan is shared with the PCP. The MLTSS enrollee population includes individuals with complex health and social support needs. Given their complex needs, they require high levels of care coordination. Effective care coordination for complex populations, such as MLTSS enrollees, includes developing and sharing individualized care plans, which are associated with numerous positive health outcomes.

Numerator Statement: Medicaid MLTSS enrollees who have a care plan (or part of a care plan) that was transmitted to their PCP within 30 days of the care plan's development or update date.

Denominator Statement: Medicaid MLTSS enrollees age 18 years and older who had a care plan developed or updated in the measurement year.

Denominator Exclusions: Exclude enrollees in the denominator who were not enrolled in an MLTSS plan for at least 30 days after a care plan's development or update date.

Exclude enrollees who refuse to have their care plan shared.

Measure Type: Process

Data Source: Management Data, Other, Paper Medical Records

Level of Analysis: Health Plan

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

3319: LTSS Comprehensive Assessment and Update

3324: LTSS Comprehensive Care Plan and Update

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? This measure is grouped with two other measures that assess the continuum of assessment, care

planning and care coordination. This continuum of care is described in greater detail in the accompanying Evidence Attachment form.

- LTSS Comprehensive Assessment and Update
- LTSS Comprehensive Care Plan and Update

Staff Preliminary Analysis: New Measure

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☐ **Yes** ☒ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☐ **Yes** ☒ **No**
- **Evidence graded?** ☐ **Yes** ☒ **No**

Evidence Summary

This measure would address the lack of standardization by assessing the percentage of MLTSS beneficiaries for whom all or part of the care plan was transmitted to PCPs, and the number of days between when the care plan was first developed or updated and then shared.

- The developer provides a logic model describing the steps between the process of completing a comprehensive assessment and care plan and the outcome of improvement in quality of life.
- There is no systematic review of studies of care planning in MLTSS programs. The developer conducted a targeted literature review to gather evidence in support of the measure.
    - There is evidence that PCPs who do not consistently receive communication from specialists were significantly more likely to report that their ability to provide high quality care was jeopardized (O'Malley & Reschovsky, 2011)
    - Systematic review of HBPC program evaluations found that among the nine studies that met high evidence standards, eight resulted in substantial reductions in at least one of the outcomes (emergency department visits, hospitalizations, hospital beds days of care, long-term care admissions, and long-term care bed days of care) with seven demonstrating reductions in at least two of these outcomes (Stall et al., 2014)
    - An ongoing evaluation of the Independence at Home demonstration, a HBPC program sponsored by CMS, found that during the second year of the program, all 15 participating practices improved performance on at least two of the six quality measures, and four practices met the performance thresholds for all six quality measures, including: annual documentation of patient preferences; all-cause hospital readmissions within 30 days; and avoidable hospital admissions and emergency room visits (CMS, 2017)
- Technical Expert Panel (TEP) convened in 2013 to provide insight into the priority areas for measurement and the usefulness and feasibility of the identified measures for MLTSS plans. The TEP was comprised of individuals representing multiple perspectives from the MLTSS community, including consumers, practitioners, health

plans, the federal government, and state governments. A second TEP was convened in 2015 with a similar composition to inform the continued development and testing of this measure.

Exception to evidence
n/a

Questions for the Committee:

- *What is the relationship of this measure to patient outcomes?*
- *How strong is the evidence for this relationship?*
- *Is the evidence directly applicable to the process of care being measured?*
- *Is there evidence of a systematic assessment of expert opinion beyond those involved in developing the measure?*

Guidance from the Evidence Algorithm

Process measure not based on systematic review (Box3) -> Empirical evidence without SR or grading (Box 7) -> Empirical evidence includes all studies in the body of evidence (Box 8) -> High-moderate quality of evidence (box 9) -> Moderate

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE: N/A

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

The measure addresses the lack of coordination between LTSS and medical care providers by ensuring a patient's care plan is shared with the PCP. Performance data provided is from five MLTSS health plans representing at least two or more of the major LTSS sub-populations. The data demonstrating the proportion of beneficiaries who have a care plan shared with their PCP shows significant room for improvement:

Percent of beneficiaries with care plan shared with PCP

Mean 6.5%

Standard Deviation 10.2%

Minimum: 0%

Maximum 23.4%

Developer provided additional performance data on care plan sharing rates by enrollee type that shows approximately 30 percent of the enrollees in the measure's denominator (133 enrollees) had their care plan shared with a PCP or a key LTSS provider at least once in the measurement period (see below). Among all the 133 enrollees who had an initial care plan or care plan update shared, 63 percent were shared within 30 days, and most were shared within 10 days.

Table 1. Care plan sharing rates by enrollee type

Enrollees with	**Frequency**	**Percentage of enrollees with a care plan eligible for measure (n =438)**
Documentation of a care plan shared with an eligible provider (PCP or Key LTSS)	**133**	**30.4**
Care plan shared within 30 days from its creation	84	19.2
Care plan shared after 30 days of its creation	21	4.8
Data entry error*	28	6.4
No documentation of a care plan shared with an eligible provider	**305**	**69.6**

Source: Mathematica analysis of enrollee data from 5 MLTSS plans.

The developer cites additional literature indicating an opportunity for improvement:

- Coordination, when it occurs, is idiosyncratic and often depends on the efforts of the care coordinator to communicate with all relevant parties and to arrange for information to flow (Saucier & Burwell, 2015).
- Confusion regarding regulations protecting patient health information can often hinder necessary information exchange (McGinn-Shapiro et al., 2015)
- Technology has been shown a critical barrier to coordination between LTSS and medical care providers (NCQA, 2015)

Disparities

- The developer did not find any research on disparities in the sharing of care plans among the LTSS enrollee population, however studies have identified persistent racial and ethnic disparities regarding advanced care planning (Barwise et al., 2016; Effiong & Myrick, 2012; Garrido et al., 2014)
- The Congressional Budget Office identified racial and ethnic disparities in the need for LTSS. More specifically, it found that older black and Hispanic individuals have higher rates of functional impairment than whites (CBO, 2013).
- The developer collected information about race and ethnicity during testing. However, due to the overall low rates, they do not believe additional analysis of disparities would provide meaningful information.

Questions for the Committee:

- *Is there a gap in care that warrants a national performance measure?*
- *Are you aware of evidence that disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments: Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence

Comments

**This is an important intermediate process measure. Community providers (home care, ALF, SNF, etc.) are critical care providers and linking to the PCP is essential; however, there are no existing measures to assess. The real outcome measure is that the PCP, with this knowledge, improves the patient's care - but this measure facilitates the ability to get to that phase. The literature clearly makes a case for importance and relevance and the measures from the five plans establishes that this is possible and that best practice evidence may very well exist from which to make system improvements. Suggestions: we do need clear descriptors of the setting/providers that are included AND what is included in the care plan measure. The use of technology in the collection seems to vary - concern this may alter the measure (?). Expanding beyond Medicaid may quickly become a consideration as other populations struggle with the coordination of care challenges. I do wonder with the advancement of HIE, is there a simpler way to implement this through that existing structure. And what is the burden to providers – this may be lightened via a HIE.

** "How does the evidence relate to the specific structure, process, or outcome being measured? No. Virtually all of the references provided did not reflect the specific intent of this measure and were used to extrapolate potential benefit without explicitly demonstrated linkages.

Does it apply directly or is it tangential? The evidence presented is by in large tangential at best across most of the references provided for this measure. There is an implicit assumption that improved communication is linked to tangible utilization and clinical outcomes, but that assumption is largely unproven or weak at best from the evidence provided.

How does the structure, process, or outcome relate to desired outcomes? Other than meeting administrative requirements for enrollment in government programs, no discrete health outcomes are specified in the evidence analysis provided for this measure other than hypothetical expectations.

For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? N/A"

**There appears to be indirect, but not direct evidence, in support of this process measure's association with outcomes.

**Empirical data are provided and are relatively complete. Some evidence is tangential (for example, data that relates PCP - specialist communication.) Overall, this is a process measure with moderate empirical evidence relating proposed process to outcomes. I am not aware of any other data that should be included.

1b. Performance Gap

Comments

**Yes, preliminary performance data was measures with five MLTSS health plans. Receipt rate very low, supporting need to understand this measure. Full data included. One health plan demonstrated a meaningful increase in scores - worthy of further understanding. Disparity information not included and no literature to support need to do so.

**There were no specific data presented documenting the performance gap in the evidence review, just an acknowledgment that states are inconsistent and highly variable in their requirements regarding sharing of LTSS related plans of care.

Performance data provided is from five MLTSS health plans representing at least two or more of the major LTSS sub-populations. The data demonstrating the proportion of beneficiaries who have a care plan shared with their PCP shows significant room for improvement:

Percent of beneficiaries with care plan shared with PCP

Mean 6.5%

Standard Deviation 10.2%

Minimum: 0%

Maximum 23.4%

Nothing mentioned about healthcare disparities in these results.

No documentation of a care plan shared with an eligible provider* (*is this the same as a PCP?) 69.6%.

There is no specific outcome or follow on process identified or analyzed in the context of ""sharing"" the care plan with a PCP. I.e. nothing about ""what happens next"" with this process and also not specifically identified or pre-specified by the measure developer. "

**There was a gap identified in sharing care plans from pilot sites (there is no demographic data submitted re: disparities and demographics), reportedly because of low overall numbers.

**Performance gap in communication of care plans to PCP is demonstrated clearly. Impact (outcome) of potentially closing this gap is less clear.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications **and** Testing
2b. Validity: Testing; Exclusions**;** Risk-Adjustment**;** Meaningful Differences**;** Comparability Missing Data

2c. For composite measures: empirical analysis support composite approach

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is

precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff

Evaluation of Reliability and Validity (and composite construction, if applicable): Link A

Questions for the Committee regarding reliability:

- *Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?*
- *Is the Committee satisfied with the reliability analysis for this measures, and is a need to discuss and/or vote on reliability?*

Questions for the Committee regarding validity:

- *Do you have any concerns regarding the validity of the measure (e.g., multiple rates, exclusions, risk-adjustment approach, etc.)?*
- *Is the Committee satisfied with the validity analysis for this measures, and is a need to discuss and/or vote on validity?*

Preliminary rating for reliability:	☐ **High**	☒ **Moderate**	☐ **Low**	☐ **Insufficient**
Preliminary rating for validity:	☐ **High**	☒ **Moderate**	☐ **Low**	☐ **Insufficient**

Scientific Acceptability

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. **Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion.**

Instructions:

- Please complete this form for each measure you are evaluating.
- Please pay close attention to the skip logic directions.
- If you are unable to check a box, please highlight or shade the box for your response.
- You must answer the "overall rating" item for both Reliability and Validity. Also, be sure to answer the composite measure question at the end of the form if your measure is a composite.
- We have provided TIPS to help you answer the questions.
- We've designed this form to try to minimize the amount of writing that you have to do. That said, ***it is critical that you explain your thinking/rationale if you check boxes where we ask for an explanation*** (because this is a Word document, you can just add your explanation below the checkbox)**.** Feel free to add additional explanation, even if an explanation is not requested (but please type this underneath the appropriate checkbox).
- This form is based on Algorithms 2 and 3 in the Measure Evaluation Criteria and Guidance document (see pages 18-24). These algorithms provide guidance to help you rate the Reliability and Validity subcriteria. ***We ask that you refer to this document when you are evaluating your measures***.
- Please contact Methods Panel staff if you have questions (methodspanel@qualityforum.org).

Measure Number: 3325

Measure Title: LTSS Shared Care Plan with Primary Care Practitioner

RELIABILITY

1. Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?
   NOTE: NQF staff will conduct a separate, more technical, check of eMeasure (eCQM) specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.

   TIPS: Consider the following: Are all the data elements clearly defined? Are all appropriate codes included? Is the logic or calculation algorithm clear? Is it likely this measure can be consistently implemented?

   ☒Yes (go to Question #2)

   ☐No (please explain below, and go to Question #2) NOTE that even though ***non-precise specifications should result in an overall LOW rating for reliability***, we still want you to look at the testing results.

2. Was empirical reliability testing (at the data element or measure score level) conducted using statistical tests with the measure as specified?

   TIPS: Check the 2nd "NO" box below if: only descriptive statistics provided; only describes process for data management/cleaning/computer programming; testing does not match measure specifications (i.e. data, eMeasure, level of analysis, patients)

   ☒Yes (go to Question #4)

   ☐No, there is reliability testing information, but *not* using statistical tests and/or not for the measure as specified OR there is no reliability testing (please explain below then go to Question #3)

3. Was **empirical VALIDITY testing** of patient-level data conducted?

   ☐Yes (use your rating from data element validity testing – Question #16- under Validity Section)

   ☐No (please explain below and rate Question #11: OVERALL RELIABILITY as INSUFFICIENT and proceed to the VALIDITY SECTION)

4. Was reliability testing conducted with computed performance measure scores for each measured entity?

   TIPS: Answer no if: only one overall score for all patients in sample used for testing patient-level data

   ☒Yes (go to Question #5)

   ☐No (go to Question #8)

5. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? *NOTE: If multiple methods used, at least one must be appropriate.*

   TIPS: Examples of appropriate methods include signal-to-noise analysis (e.g. Adams/RAND tutorial); random split-half correlation; other accepted method with description of how it assesses reliability of the performance score.

   ☒Yes (go to Question #6) Split sample reliability was assessed using ICC

   ☐No (please explain below then go to Question #8)

6. **RATING (score level)** - What is the level of certainty or confidence that the performance measure scores are reliable?

   TIPS: Consider the following: Is the test sample adequate to generalize for widespread implementation? Do the results demonstrate sufficient reliability so that differences in performance can be identified?

   ☒High (go to Question #8)

☐Moderate (go to Question #8)

☐Low (please explain below then go to Question #7)

7. Was other reliability testing reported?

   ☐Yes (go to Question #8)

   ☐No (rate Question #11: OVERALL RELIABILITY as LOW and proceed to the VALIDITY SECTION)

8. Was reliability testing conducted with patient-level data elements that are used to construct the performance measure?

   TIPS: Prior reliability studies of the same data elements may be submitted; if comparing abstraction to "authoritative source/gold standard" see Validity Section Question #15)

   ☒Yes (go to Question #9)

   ☐No (if there is score-level testing, rate Question #11: OVERALL RELIABILITY based on score-level rating from Question #6; otherwise, rate Question #11: OVERALL RELIABILITY as INSUFFICIENT. Then proceed to the VALIDITY SECTION)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   TIPS: For example: inter-abstractor agreement (ICC, Kappa); other accepted method with description of how it assesses reliability of the data elements

   Answer no if: only assessed percent agreement; did not assess separately for all data elements (at least numerator, denominator, exclusions)

   ☒Yes (go to Question #10) Cohen's kappa statistic used to evaluate IRR

   ☐No (if no, please explain below and rate Question #10 as INSUFFICIENT)

10. **RATING (data element)** – Based on the reliability statistic and scope of testing (number and representativeness of patients and entities), what is the level of certainty or confidence that the data used in the measure are reliable?

    TIPS: Consider the following: Is the test sample adequate to generalize for widespread implementation? Can data elements be collected consistently?

    ☐Moderate (if score-level testing was NOT conducted, rate Question #11: OVERALL RELIABILITY as MODERATE)

    ☒Low (if score-level testing was NOT conducted, rate Question #11: OVERALL RELIABILITY as LOW)

    ☐Insufficient (go to Question #11)

    Elements in the LTSS Shared Care Plan with Primary Care Practitioner measure were assessed too infrequently among the 144 paired assessments (<30) to allow for inter-rater reliability analysis. The data element indicating the care plan shared met the threshold for slight reliability.

11. OVERALL RELIABILITY RATING

OVERALL RATING OF RELIABILITY taking into account precision of specifications and all testing results:

☐High (NOTE: Can be HIGH only if score-level testing has been conducted)

☒Moderate (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

☐Low (please explain below) [NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete]

☐Insufficient (please explain below) [NOTE: For most measure types, testing at both the score level and the data element level is not required]

VALIDITY

ASSESSMENT OF THREATS TO VALIDITY

1. Were all potential threats to validity that are relevant to the measure empirically assessed?

   TIPS: Threats to validity include: exclusions; need for risk adjustment; Able to identify statistically significant and meaningful differences; multiple sets of specifications; missing data/nonresponse.

   ☒Yes (go to Question #2)

   ☐No (please explain below and go to Question #2) [NOTE that even if ***non-assessment of applicable threats should result in an overall INSUFFICENT rating for validity***, we still want you to look at the testing results]

2. Analysis of potential threats to validity: Any concerns with measure exclusions?

   TIPS: Consider the following: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure? Are the exclusions/exceptions of sufficient frequency and variation across providers to be needed (and outweigh the data collection burden)? If patient preference (e.g., informed decisionmaking) is a basis for exclusion, does it impact performance and if yes, is the measure specified so that the information about patient preference and the effect on the measure is transparent?

   ☐Yes (please explain below then go to Question #3)

   ☒No (go to Question #3)

   ☐Not applicable (i.e., there are no exclusions specified for the measure; go to Question #3)

3. Analysis of potential threats to validity: Risk-adjustment (applies to all outcome, cost, and resource use measures; may also apply to other types of measure)

   ☒Not applicable (e.g., structure or process measure that is not risk-adjusted; go to Question #4)

   a. Is a conceptual rationale for social risk factors included? ☐Yes ☐No

   b. Are social risk factors included in risk model? ☐Yes ☐No

   c. Any concerns regarding the risk-adjustment approach?

   *TIPS: Consider the following: If a justification for **not risk adjusting** is provided, is there any evidence that contradicts the developer's rationale and analysis? If the developer asserts there is **no conceptual basis** for adjusting this measure for social risk factors, do you agree with the rationale? **If risk adjusted**: Are the candidate and final variables included in the risk adjustment model adequately described for the measure to be implemented? Are the candidate and final variables included in the risk adjustment model adequately described for the measure to be implemented? Are all of the risk adjustment variables present at the start of care (if not, do you agree with the rationale)? If social risk factors are not included in the risk-adjustment approach, do you agree with the developer's decision? Is an appropriate risk-adjustment strategy included in the measure (e.g., adequate model discrimination and calibration)? Are all statistical model specifications included, including a "clinical model only" if social risk factors are included in the final model?*

   ☐Yes (please explain below then go to Question #4)

   ☐No (go to Question #4)

4. Analysis of potential threats to validity: Any concerns regarding ability to identify meaningful differences in performance or overall poor performance?

   ☒Yes (please explain below then go to Question #5)

   ☐No (go to Question #5)

   Very poor overall performance observed by developer. However, developer indicates that health plan 01 had rates demonstrating a statistically significant difference from the mean. This finding indicates that the measure as specified does have the potential to identify meaningful differences in performance between health plans.

5. Analysis of potential threats to validity: Any concerns regarding comparability of results if multiple data sources or methods are specified?

   ☐Yes (please explain below then go to Question #6)

   ☐No (go to Question #6)

   ☒Not applicable (go to Question #6)

The developer did not provide an analysis of the comparability of results.

6. Analysis of potential threats to validity: Any concerns regarding missing data?

   ☐Yes (please explain below then go to Question #7)

   ☒No (go to Question #7)

ASSESSMENT OF MEASURE TESTING

7. Was empirical validity testing conducted using the measure as specified and appropriate statistical test?

   Answer no if: face validity; only refer to clinical evidence; only descriptive statistics; only describe process for data management/cleaning/computer programming; testing does not match measure specifications (i.e. data, eMeasure, level, setting, patients).

   ☐Yes (go to Question #10) [NOTE: If appropriate empirical testing has been conducted, then evaluation of face validity is not necessary. Go to Question #8 **only if** there is insufficient information provided to evaluate data element and score-level testing.]

   ☒No (please explain below then go to Question #8)

   Score level empirical testing was done, but results were inconclusive (e.g. neither validated, nor invalidated the measure). These may be viewed in the testing submission.

8. Was face validity systematically assessed by recognized experts to determine agreement on whether the computed performance measure score from the measure as specified can be used to distinguish good and poor quality?

   TIPS: Answer no if: focused on data element accuracy/availability/feasibility/other topics; the degree of consensus and any areas of disagreement not provided/discussed.

   ☒Yes (go to Question #9)

   ☐No (please explain below and rate Question #17: OVERALL VALIDITY as INSUFFICIENT)

9. **RATING (face validity)** - Do the face validity testing results indicate substantial agreement that the performance measure score from the measure as specified can be used to distinguish quality AND potential threats to validity are not a problem, OR are adequately addressed so results are not biased?

   ☒Yes (if a NEW measure, rate Question #17: OVERALL VALIDITY as MODERATE)

   ☐ Yes (if a MAINTENANCE measure, do you agree with the justification for not conducting empirical testing? If no, rate Question #17: OVERALL VALIDITY as INSUFFICIENT; otherwise, rate Question #17: OVERALL VALIDITY as MODERATE)

   ☐No (please explain below and rate Question #17: OVERALL VALIDITY AS LOW)

Response	Question 5: Would high performance on this measure indicate that a health plan is providing higher quality care?	Question 6: In the future, do you think that performance scores on this measure will distinguish between good and poor performance?
Strongly Agree	1	1
Agree	7	6
Disagree	2	3
Strongly Disagree	3	3
No response	0	0
Total % Agree	62%	54%

See Testing Appendix for summaries of written feedback on Systematic Assessment of Face Validity

Question 5:

- Commenters noted that the measure was specified as a process measure and does not correlate patient outcomes to a care plan.
- TEP members that supported the measure were in agreement that sharing a care plan with the PCP is a first step in improving performance.

Question 6:

- One TEP member suggested extending the 30-day transmission window to 45 days.
- One TEP member noted that some percentage of enrollees in managed LTSS plans do not have a primary care physician.
- One TEP member suggested that there was greater importance for the patient or consumer to receive a copy of the care plan than the PCP

10. Was validity testing conducted with computed performance measure scores for each measured entity?

    TIPS: Answer no if: one overall score for all patients in sample used for testing patient-level data.

    ☐Yes (go to Question #11)

    ☐No (please explain below and go to Question #13)

11. Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?

    TIPS: For example: correlation of the performance measure score on this measure and other performance measures; differences in performance scores between groups known to differ on quality; other accepted method with description of how it assesses validity of the performance score

    ☐Yes (go to Question #12)

    ☐No (please explain below, rate Question #12 as INSUFFICIENT and then go to Question #14)

12. **RATING (measure score)** - Based on the measure score results (significance, strength) and scope of testing (number of measured entities and representativeness) and analysis of potential threats, what is the level of certainty or confidence that the performance measure scores are a valid indicator of quality?

    ☐High (go to Question #14)

    ☐Moderate (go to Question #14)

    ☐Low (please explain below then go to Question #13)

☐Insufficient

13. Was other validity testing reported?

    ☐Yes (go to Question #14)

    ☐No (please explain below and rate Question #17: OVERALL VALIDITY as LOW)

14. Was validity testing conducted with patient-level data elements?

    TIPS: Prior validity studies of the same data elements may be submitted

    ☐Yes (go to Question #15)

    ☐No (please explain below and rate Question #17: OVERALL VALIDITY as INSUFFICIENT if no score-level testing was conducted, otherwise, rate Question #17: OVERALL VALIDITY based on score-level rating from Question #12)

    Systematic assessment of face validity surveyed 13 member technical expert panel.

15. Was the method described and appropriate for assessing the accuracy of ALL critical data elements? *NOTE that data element validation from the literature is acceptable.*

    TIPS: For example: Data validity/accuracy as compared to authoritative source- sensitivity, specificity, PPV, NPV; other accepted method with description of how it assesses validity of the data elements.

    Answer No if: only assessed percent agreement; did not assess separately for all data elements (at least numerator, denominator, exclusions)

    ☐Yes (go to Question #16)

    ☐No (please explain below and rate Question #16 as INSUFFICIENT)

16. **RATING (data element)** - Based on the data element testing results (significance, strength) and scope of testing (number and representativeness of patients and entities) and analysis of potential threats, what is the level of certainty or confidence that the data used in the measure are valid?

    ☐Moderate (if score-level testing was NOT conducted, rate Question #17: OVERALL VALIDITY as MODERATE)

    ☐Low (please explain below) (if score-level testing was NOT conducted, rate Question #17: OVERALL VALIDITY as LOW)

    ☐Insufficient (go to Question #17)

17. OVERALL VALIDITY RATING

OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.

☐High (NOTE: Can be HIGH only if score-level testing has been conducted)

☒Moderate (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐Low (please explain below) [NOTE: Should rate LOW if you believe that there are threats to validity and/or threats to validity were not assessed]

☐Insufficient (if insufficient, please explain below) [NOTE: For most measure types, testing at both the score level and the data element level is not required] [NOTE: If rating is INSUFFICIENT for all empirical testing, then go back to Question #8 and evaluate any face validity that was conducted, then reconsider this overall rating.]

Empirical validity provided was inconclusive (neither validated, nor invalidated the measure). Rating was based on face validity results (which are lower than desired -- most likely due to the extremely low performance rate of the plans in the testing sample).

Committee Pre-evaluation Comments: Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability- Specifications

Comments

**Yes. Very through. Outstanding questions: clear descriptors on who are included as providers; what is included in the "care plan" such as meds, risk factors, etc.; does it include contact information. Under Table 3 - do question the impact of the reliability measure with date of first disenrollment -- not clear on the process but concern this impacts the integrity of the data. Details on provider types will be warranted (only those paid by MA in the community of non-hospital facility?)

**Elements in the LTSS Shared Care Plan with Primary Care Practitioner measure were assessed too infrequently among the 144 paired assessments (<30) to allow for inter-rater reliability analysis. The data element indicating the care plan shared met the threshold for slight reliability. Disagree with preliminary rating of "moderate" would put this as "low" to "insufficient". Noted in the Feasibility section is also this information:

• Data elements used to calculate this measure are abstracted from record by someone other than the person obtaining original information.

• This measure is primarily collected from health plan and case management records, many of which are electronic.

o Some data elements are in defined fields in electronic sources

o The data elements needed for this measure are not standardized (i.e. may be in free text rather than structured fields).

"

**limited information submitted

**Specifications seem clear, but testing samples were small.

2a2. Reliability- Testing

Comments

**see above

need a better understanding of the ""other elements"" that were assessed too infrequently to allow IRR analysis.

**Yes, believe it is low to insufficient per comments above. Disagree that it is "moderate".

**What constitutes a "care plan" might be inconsistently applied because the definition has some inherent ambiguity. Same with what constitutes an "update".

**Inadequate data for inter-rater reliability analysis as noted in the ""Scientific Acceptability"" document. Reliability seems acceptable with regard to other elements."

2b1. Validity—Testing

2b4-7. Threats to Validity

2b4. Meaningful Differences

Comments

**Tested both performance and face validity. Panels large (clarification - that is the group included in the packet - correct?) and inclusive; however, do not see as many "providers" as I would have thought as this would be the group who would best understand validity. I did not see the TEP comments in my packet (?) so cannot fully describe fully assess response to comments. Missing data not fully described, but was discussed. Did compare with other care plan recommended measure rates which was helpful. Overall TEP recommended measure. Table 6 - row calculations would be helpful; clear that >80% agree. More discussion on harmonizing is warranted in moving forward. Patients can refuse to have their care plans shared which is confusing since this is a managed care program; curious how often this actually occurs. Any other exclusion trends?

**Do you have any concerns with the testing results?

2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) Yes--insufficient data to generate statistically valid inter-rater reliability.

2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? Not demonstrated other than low initial response rates. Lack of response could be directly related to burden of meeting expectations of the measure developers.

2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? Not done.

2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure? Unsure.

**Transmission of a care plan does not reflect receipt of the care plan by the PCP or quality of the care plan itself (both of which are likely more important). Lack of direct evidence to link success in this process measure to improved outcomes is a barrier to validity.

**Some empirical testing was performed but was inconclusive.

Measure validity relies heavily on Technical Expert Panel for face validity assessment. Expert panel agreement regarding face validity is tempered, at best (62% and 54% agree on questions 5 and 6 in the Systematic Assessment of Face Validity). If one more expert disagreed with either question, face validity questions would be 50% or less. Is a simple majority (7 or 8 out of of 13 experts) enough to support face validity? Seems like that should be much higher.

2b2-3. Other Threats to Validity
2b2. Exclusions
2b3. Risk Adjustment
Comments
**n/a None of these threats were explicitly evaluated or mentioned in the analysis.
**Exclusions are appropriate. No risk adjustment needed.

Criterion 3. Feasibility

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- Data elements used to calculate this measure are abstracted from record by someone other than the person obtaining original information
- This measure is primarily collected from health plan and case management records, many of which are electronic.
    - Some data elements are in defined fields in electronic sources
    - The data elements needed for this measure are not standardized (i.e. may be in free text rather than structured fields).
- The developer noted that as the LTSS fields move forward, more of the data elements used in the measure will become structured fields.
- No fees or licensing are currently required.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Does the Committee agree that measurement in this area will drive standardization?*

Preliminary rating for feasibility: ☐ **High** ☐ **Moderate** ☒ **Low** ☐ **Insufficient**

RATIONALE: Data elements needed for this measure are not currently standardized.

Committee Pre-evaluation Comments: Criteria 3: Feasibility

3. Feasibility
Comments
**A consistent approach will need to be well-defined as the data sources are originating in such a variety of setting (home care, SNFs) with small and large operations. Use of a HIE might help in a consistent mechanism to deliver to the

PCP - concern if some information comes via EHR, fax, mail, email, it may make it difficult for the PCP to manage. Agree that an emeasure would be warranted for ongoing measurement
**The data is not standardized or congruent with data formatting standards making this very challenging from a measurement and evaluation standpoint. There is also no mention of the need for interoperability requirements, given that this measure evaluates a "handshake" of information and nothing more.
**Not necessarily routinely collected in a manner that makes them accessible. Not collected in a standardized way at present, but likely could be abstracted (with some effort) from the medical record kept at LTSS.
**In current environment, processes to create care plans and elements specified are not standard. Standardizing the process and elements will likely be resources intensive for health plans and potentially for providers. Unclear how much would be needed in terms of new resources to participate with and meet the measure, versus reallocation of existing resources.

Criterion 4: Usability and Use

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.
4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☐ **Yes** ☒ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☐ **Yes** ☐ **No**

Accountability program details

- This is a new measure and is not currently in use.
- This measure is intended for use by states and health plans to monitor and improve quality of care provided for the Medicaid MLTSS enrollee population.
- This measure is included in the National MLTSS Health Plan Association recommended LTSS performance measure set model.

4a.2. Feedback on the measure by those being measured or others**.** Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- N/A

Additional Feedback:

- N/A

***Questions for the Committee**:*

- *How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?*
- *How has the measure been vetted in real-world settings by those being measured or others?*

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.
4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- This is a new measure and improvement information was not provided

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- The developer reported that no unintended consequences were identified during testing

Potential harms

- N/A

Additional Feedback:

- N/A

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for Usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments: Criteria 4: Usability and Use

4a1. Use- Accountability and Transparency

Comments

**This measure is under initial endorsement review; no accountability program in place. This measure is useful as a springboard to the next phase of care coordination of the post-acute and community of care linking with the PCP to ensure goals of care are consistent and the team is working towards them together. Not sure the provider group has had the opportunity to provide feedback in the development; one of their core concerns will be ease of collection and ability to get to the PCP. Consideration will need to be given to the changes in plans and physician. The other consideration is that this is at 30 days - in some situations, this may be too late - the coordination may have needed to be considerably sooner. The real use is when the PCP and provider team coordinate care together; this measure in and of itself does not get us there, but it is a step. A soon thereafter measure will need to evaluate the quality of the information being sent (is it meaningful) and did the PCP review it to adjust the care plan.

**Accountability program details

• This is a new measure and is not currently in use.

• This measure is intended for use by states and health plans to monitor and improve quality of care provided for the Medicaid MLTSS enrollee population.

• This measure is included in the National MLTSS Health Plan Association recommended LTSS performance measure set model.

There has been no attempt to use this measure individually or in combination with other LTSS measures presented to the Committee.

The measure could be useful to improve operational efficiency in terms of assuring higher compliance/success rates from an administrative accountability standpoint.

It is very unclear from my review whether this measure can meet the standard intent of NQF, MAP and HHS to endorse ""Measures that Matter"".
**New measure, so limited info. It appears that CMS will ask states / plans to report on this measure.
**Given that this measure would likely require additional (or reallocated) HP resources to achieve success, it would be nice to see a stronger tie to outcome. In particular, this measure is inferring a patient outcome benefit related specifically to sharing HP care plan with PCP. The measure has (at best) moderate face validity by TEP, lacks direct evidence of outcome benefit, and may consume additional resources. Based on data presented, it is not clear that this measure meets goals of improving efficiency of care or reliably distinguishing higher quality or better performing by a health plan.

Criterion 5: Related and Competing Measures

Related or competing measures

Related measures include:

- 3319 LTSS Comprehensive Assessment and Update
- 3324 LTSS Comprehensive Care Plan and Update
- 3326 LTSS Re-Assessment/Care Plan Update after Inpatient Discharge

Harmonization

N/A

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

N/A

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: January 18, 2018

Comments:

Morgan Buchko, Meridian Health Plan

This requires the plan to track providing an updated or new care plan to the PCP within 30 days. If we are going to be required to report on this, we will need a spec around what constitutes a significant change that requires the PCP notificatoin.

- **Zero NQF members have submitted a support/non-support choice.**

Developer Submission

Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

Brief Measure Information

NQF #: 3325

Corresponding Measures:

De.2. Measure Title: Long Term Services and Supports (LTSS) Shared Care Plan with Primary Care Practitioner

Co.1.1. Measure Steward: Centers for Medicare & Medicaid Services, Centers for Medicaid & CHIP Services

De.3. Brief Description of Measure: This measure assesses the percentage of Medicaid Managed Long Term Services and Supports (MLTSS) Plan enrollees with a care plan for whom all or part of the care plan was transmitted to the primary care practitioner (PCP) within 30 days of the care plan's development or update.

1b.1. Developer Rationale: This measure would address the lack of coordination between LTSS and medical care providers by ensuring a patient's care plan is shared with the PCP. The MLTSS enrollee population includes individuals with complex health and social support needs. Given their complex needs, they require high levels of care coordination. Effective care coordination for complex populations, such as MLTSS enrollees, includes developing and sharing individualized care plans, which are associated with numerous positive health outcomes.

S.4. Numerator Statement: Medicaid MLTSS enrollees who have a care plan (or part of a care plan) that was transmitted to their PCP within 30 days of the care plan's development or update date.

S.6. Denominator Statement: Medicaid MLTSS enrollees age 18 years and older who had a care plan developed or updated in the measurement year.

S.8. Denominator Exclusions: Exclude enrollees in the denominator who were not enrolled in an MLTSS plan for at least 30 days after a care plan's development or update date.

Exclude enrollees who refuse to have their care plan shared.

De.1. Measure Type: Process

S.17. Data Source: Management Data, Other, Paper Medical Records

S.20. Level of Analysis: Health Plan

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

3320:LTSS Comphrensive Assessment, Care Planning, and Coordination

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? This measure is grouped with two other measures that assess the continuum of assessment, care planning and care coordination. This continuum of care is described in greater detail in the accompanying Evidence Attachment form.

- LTSS Comprehensive Assessment and Update
- LTSS Comprehensive Care Plan and Update

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

LTSS_Shared_Care_Plan_Evidence_Attachment.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission? Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

1a Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** TBD

Measure Title: LTSS Shared Care Plan with Primary Care Practitioner

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: N/A

Date of Submission:

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.

- Efficiency: [6] evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome:

☐Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (*A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)*

☐ Intermediate clinical outcome (*e.g., lab value*):

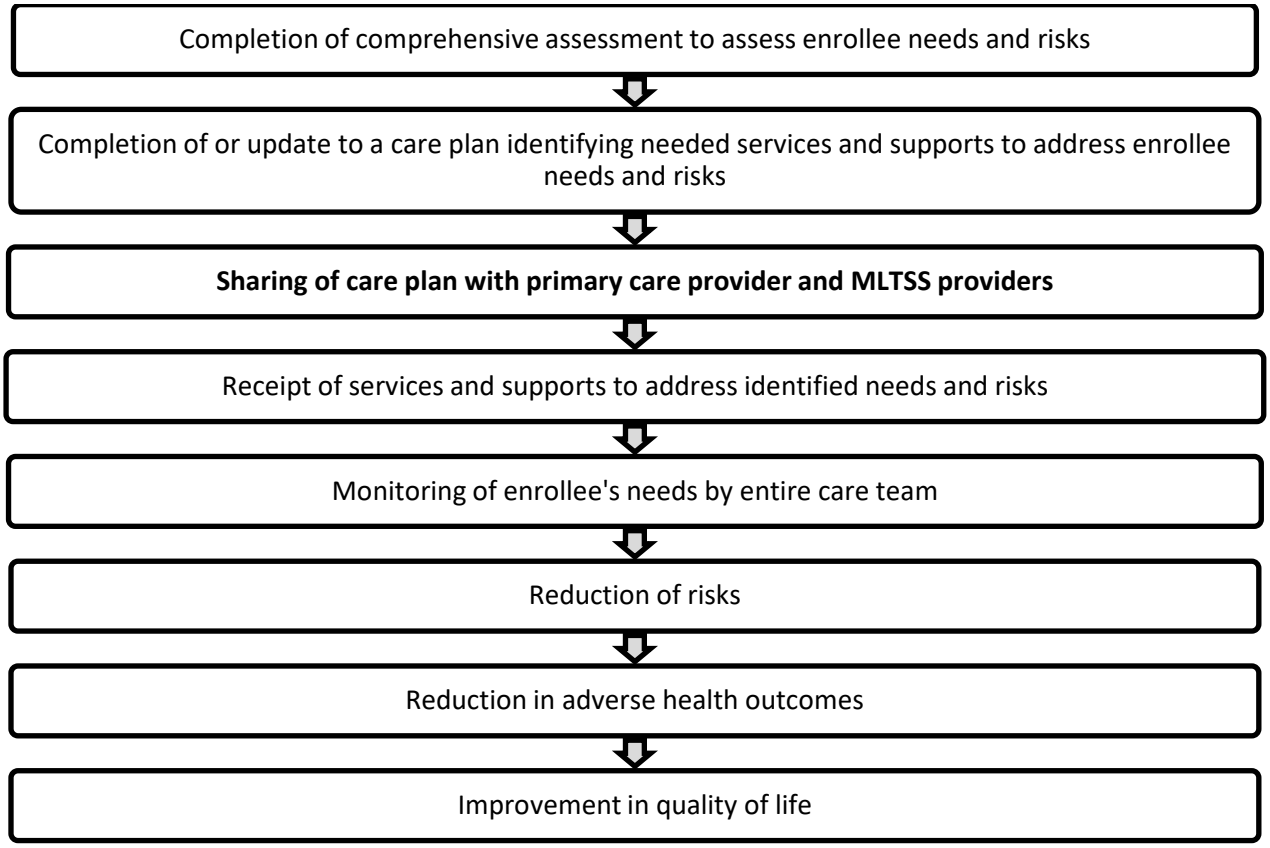
☒ Process: The percentage of Medicaid Managed Long Term Services and Supports (MLTSS) Plan enrollees with a care plan for whom all or part of the care plan was transmitted to the primary care practitioner (PCP) within 30 days of the care plan's development or update.

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.



1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

Not applicable.

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

Not applicable. Not an outcome measure.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐Clinical Practice Guideline recommendation (with evidence review)

☐US Preventive Services Task Force Recommendation

☐Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐Other

Not applicable. Evidence is not based on a systematic review.

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

In the absence of a systematic review of studies of shared care plans in MLTSS programs, the project team conducted a targeted literature review to gather evidence in support of this measure. We searched for academic journal articles, gray literature, and federal and state agency reports published in the last 23 years using PubMed, Google, and Google Scholar.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

The Medicaid MLTSS enrollee population includes individuals with complex health and social support needs, such as individuals with physical, cognitive, and mental disabilities and older adults with multiple functional limitations and chronic conditions (MACPAC, 2016; KFF, 2015). Given their complex needs, Medicaid MLTSS enrollees often require a wide range of services and supports and high levels of care coordination (Saucier & Burwell, 2015). Delivering effective care coordination for complex populations, such as Medicaid MLTSS enrollees, begins with conducting and regularly updating a comprehensive assessment to identify enrollees' needs, developing and regularly updating an individualized care plan to indicate the specific services and supports to be provided, and sharing care plans to inform care team

members of services that should be coordinated (Rich et al., 2012). State Medicaid agencies have implemented numerous Medicaid MLTSS care coordination models (Saucier & Burwell, 2015). Most models require the development of a care plan at initial enrollment and on a regular basis thereafter, as well as the use of team based care to implement the care plan. Similarly, numerous other programs that deliver care to individuals who are similar to (or in some cases the same as) Medicaid MLTSS enrollees require care plans and team based care, including patient-centered medical homes, Medicare managed care plans (e.g., Special Needs Plans, Financial Alignment Initiative dual eligible enrollee demonstration plans), state Medicaid home and community-based services 1915(c) waiver programs, the Program for All-Inclusive Care for the Elderly (PACE), and Medicaid Health Homes. In order for team-based care to be effective, providers must share the care plan and communicate changes and updates to the care plan so that all members of the care team have a complete picture of the person's needs, preferences, and services and supports provided.

Well-developed care plans are associated with numerous positive outcomes, including improving patient-provider and provider-provider communication, encouraging care team accountability, flagging potential concerns for future evaluation, and promoting individuals' and caregivers' self-management (Rich et al., 2012). Documenting goals alone has been linked to numerous positive health outcomes across different care settings, such as greater improvements in health and functioning, in a variety of MLTSS-related populations, such as those with dementia (Clare et al., 2015), coronary heart disease (Janssen et al., 2013), stroke (Warner et al., 2015), end stage renal disease (Kauric-Klein, 2012), and rehabilitation needs (Muller et al., 2011).

In all of the above studies, the care plan was shared with the clinicians providing care. The sharing of information between providers (both MLTSS and medical care providers) is a critical component of providing coordinated person-centered care and breaking down the silos that exist between medical care and MLTSS providers.

Evidence to Support Sharing of Information with Primary Care Providers (PCP)

There is no direct evidence of the impact of sharing MLTSS care plan information with the PCP on outcomes. However, there is related evidence demonstrating the importance of sharing information about other types of specialty care with PCPs.

While primary care has been demonstrated to be associated with better health outcomes and a decrease in hospital admissions and emergency department visits, lack of communication between PCPs and specialists can hinder the effectiveness of primary care (Shi, 2012; Gandhi et al., 2000; Hanlon, 2013; O'Malley & Cunningham, 2009; O'Malley & Reschovsky, 2011). For example, 28 percent of PCPs expressed dissatisfaction with the content of information they receive from specialists and 50percent of PCPs were dissatisfied with the timeliness of information they received; within two weeks of referral visits, 40 percent of PCPs received no information from specialists, and four weeks after the referral visit, 25 percent of PCPs still had no information (Gandhi et al., 2000). Another study found 81 percent of specialists said they "always" or "most of the time" send referring PCPs notification of results and advice to patients, but only 62 percent of PCPs say they received this information (O'Malley & Reschovsky, 2011). Those PCPs who do not consistently receive communication from specialists were significantly more likely to report that their ability to provide high quality care was jeopardized.

MLTSS providers are in a unique position to provide PCPs with valuable information about an individual's risks due to their frequent presence in the patient's home. MLTSS care managers frequently conduct in-home assessments and communicate with home based care providers. They can directly observe issues such as home safety risks, potential for medication errors due to disorganized, expired or incorrect medications, food and nutrition concerns, and environmental hazards. Direct care workers, such as personal care aides, may make even more frequent home visits, sometimes daily, to provide hands-on assistance with activities of daily living such as bathing, eating and transferring, which gives them greater opportunity to observe changes in an individual's health and functional status. However, MLTSS providers may not have the authority to modify a medical care plan based on their observations. Therefore, coordination between MLTSS providers and medical care providers is critical to avoid potentially negative outcomes for individuals using MLTSS care.

Shared care planning and team-based care coordination among PCPs and home-care providers are hallmarks of home-based primary care programs (HBPC), which serve individuals with multiple chronic illnesses and functional limitations. While the frequency of communication among PCPs and other team members in the HBPC model is much greater than

that in non-HBPC programs, several studies demonstrate its effectiveness in improving quality and reducing the use of intensive care. For example, a systematic review of HBPC program evaluations found that among the nine studies that met high evidence standards, eight resulted in substantial reductions in at least one of the outcomes (emergency department visits, hospitalizations, hospital beds days of care, long-term care admissions, and long-term care bed days of care) with seven demonstrating reductions in at least two of these outcomes (Stall et al., 2014). An ongoing evaluation of the Independence at Home demonstration, a HBPC program sponsored by CMS, found that during the second year of the program, all 15 participating practices improved performance on at least two of the six quality measures, and four practices met the performance thresholds for all six quality measures, including: annual documentation of patient preferences; all-cause hospital readmissions within 30 days; and avoidable hospital admissions and emergency room visits (CMS, 2017).

Variation in the Frequency and Timeliness of Sharing of Care Plans with Providers

State MLTSS program contract provisions vary in the specificity with which they require managed care plans to facilitate sharing of care plans and clinical or other information among members' providers. For example, Massachusetts' Senior Care Options (SCO) program requires that the plans "ensure linkages among the PCP, the PCT [primary care team], and any appropriate acute, long-term care, or behavioral health providers to keep all parties informed about utilization of services," and develop protocols for sharing clinical and individualized plan of care information among the enrollee's caregivers. However, most other state MLTSS contract language is more general with regard to coordination among medical care and MLTSS providers and most are silent with respect to sharing members' care plans among providers (Rivard et al., 2013).

This measure would address the lack of standardization by assessing the percentage of MLTSS beneficiaries for whom all or part of the care plan was transmitted to PCPs, and the number of days between when the care plan was first developed or updated and then shared.

1a.4.2 What process was used to identify the evidence?

In the absence of a systematic review of studies of shared care plans in MMLTSS programs, the project team conducted a targeted literature review to gather evidence in support of this measure. We searched for academic journal articles, gray literature, and federal and state agency reports published in the last 23 years using PubMed, Google, and Google Scholar. We also convened a technical expert panel (TEP) in 2013 to provide insight into the priority areas for measurement and the usefulness and feasibility of the identified measures for MLTSS plans. The 2013 TEP was comprised of individuals representing multiple perspectives from the MLTSS community including consumers, practitioners, health plans, the federal government, and state governments. Under the current contract in 2016, we convened a new TEP (21 members) with a similar composition to inform the continued development and testing of this measure.

We also built upon an environmental scan of Assessment and Care Planning measures conducted under a previous CMS contract (Contract No. HHSM-500-2010-00026I/HHSM-500-T0011).

1a.4.3. Provide the citation(s) for the evidence.

CMS (2017). "Independence at Home Demonstration, Corrected Performance Year 2 Results." https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2017-Fact-Sheet-items/2017-01-19.html

Gandhi, T. K., Sittig, D. F., Franklin, M., Sussman, A. J., Fairchild, D. G., & Bates, D. W. (2000). Communication breakdown in the outpatient referral process. *Journal of General Internal Medicine, 15*(9), 626-631.

Hanlon, C. (2013). Measuring and Improving Care Coordination: Lessons from ABCD III. Portland, ME: The National Academy for State Health Policy.

Janssen, V., De Gucht, V., Dusseldorp, E., & Maes, S. (2013). Lifestyle modification programmes for patients with coronary heart disease: a systematic review and meta-analysis of randomized controlled trials. *European Journal of Preventive Cardiology, 20*(4), 620-640.

Kaiser Family Foundation (KFF). (2015). Medicaid and Long-Term Services and Supports: A Primer. Available at http://kff.org/medicaid/report/medicaid-and-long-term-services-and-supports-a-primer/.

Kauric-Klein, Z. (2012). Improving blood pressure control in end stage renal disease through a supportive educative nursing intervention. *Nephrology Nursing Journal, 39*(3), 217-228.

Medicaid and CHIP Payment and Access Commission (MACPAC). (2016). Users of long-term services and supports. Available at https://www.macpac.gov/subtopic/long-term-services-and-supports-population/.

Muller, M., Strobl, R., & Grill, E. (2011). Goals of patients with rehabilitation needs in acute hospitals: goal achievement is an indicator for improved functioning. *Journal of Rehabilitation Medicine, 43*(2), 145-150.

O'Malley, A. S., & Cunningham, P. J. (2009). Patient experiences with coordination of care: the benefit of continuity and primary care physician as referral source. *Journal of General Internal Medicine, 24*(2), 170-177.

O'Malley, A. S., & Reschovsky, J. D. (2011). Referral and consultation communication between primary care and specialist physicians: finding common ground. *Archives of Internal Medicine, 171*(1), 56-65.

P. Rivard, B. Jackson, J. Rachel, J. Seibert, and T. Whitworth (2013). "Environment Scan of LTSS Quality Requirements in MCO Contracts." Assistant Secretary for Planning and Evaluation (ASPE), U.S. Department of Health and Human Services. https://aspe.hhs.gov/system/files/pdf/76871/MCOcontr.pdf

Rich, E., D. Lipson, J. Libersky, and M. Parchman (2012). Coordinating Care for Adults with Complex Care Needs in the Patient-Centered Medical Home: Challenges and Solutions. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ). . Available at https://pcmh.ahrq.gov/sites/default/files/attachments/coordinating-care-for-adults-with-complex-care-needs-white-paper.pdf.

Saucier, P., & B. Burwell. (2015). Care Coordination in Managed Long-Term Services and Supports. Washington, DC: AARP Public Policy Institute. Available at http://www.aarp.org/content/dam/aarp/ppi/2015/care-coordination-in-managed-long-term-services-and-supports-report.pdf.

Shi, L. (2012). The impact of primary care: a focused review. *Scientifica (Cairo), 2012*, 432892.

Stall, N., M. Nowaczynski, and S. Sinha (2014). "Systematic Review of Outcomes from Home-Based Primary Care Programs for Homebound Older Adults." *Journal of the American Geriatrics Society,* 62(12): 2243-2251.

Warner, G., Packer, T., Villeneuve, M., Audulv, A., & Versnel, J. (2015). A systematic review of the effectiveness of stroke self-management programs for improving function and participation outcomes: self-management programs for stroke survivors. *Disability and Rehabilitation*, 1-23.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

This measure would address the lack of coordination between LTSS and medical care providers by ensuring a patient's care plan is shared with the PCP. The MLTSS enrollee population includes individuals with complex health and social support needs. Given their complex needs, they require high levels of care coordination. Effective care coordination for complex populations, such as MLTSS enrollees, includes developing and sharing individualized care plans, which are associated with numerous positive health outcomes.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample,*

characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.

These data are from five MLTSS health plans that participated in testing these measures with enrollees representing at least two or more of the major LTSS sub-populations, which include frail older adults age 65 and over, people under age 65 with physical disabilities, people with developmental or intellectual disabilities, and people with serious mental illness. All five participating plans covered Medicare benefits as well as Medicaid LTSS and had at least one of the following types of contracts: Medicare-Medicaid Plan (MMP) (integrated care), Dual-Special Needs Plan (D-SNP), or Fully Integrated Dual Eligible (FIDE) SNP. For purposes of testing, the eligible population included MLTSS members who were age 18 and older as of January 1, 2015 and enrolled on or prior to August 31, 2015, allowing for at least 120 days of enrollment during the measurement year of CY 2015.

The data demonstrating the proportion of beneficiaries who have a care plan shared with the PCP shows significant room for improvement (see results below).

Percent of beneficiaries with care plan shared with PCP

Mean 6.5%

Standard Deviation 10.2%

Minimum: 0%

Maximum 23.4%

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

There is no direct estimate for the rate of sharing care plan information between providers. However, evidence does suggest coordination and communication between LTSS providers and medical care providers is a critical gap. Among many dual eligible enrollees, which make up a large portion of LTSS enrollees, LTSS care is covered by a state Medicaid program and medical care is covered by Medicare either FFS or in a managed care arrangement. Recent research conducted in organizations providing care coordination for LTSS services found that even when financing for both Medicaid and Medicare services is integrated, care is often delivered in silos with medical and LTSS systems operating independently. One study found that establishing relationships between providers is critical for ensuring information exchange, and although technology supports such exchanges, coordinating care remains a “high touch activity.” In addition, EHRs have not been widely adopted by LTSS providers, and furthermore, existing EHRs do not incorporate the type of information needed by LTSS providers. Finally, confusion regarding regulations protecting patient health information can often hinder necessary information exchange (McGinn-Shapiro et al., 2015). Coordination, when it occurs, is idiosyncratic and often depends on the efforts of the care coordinator to communicate with all relevant parties and to arrange for information to flow (Saucier & Burwell, 2015).

Technology is one critical barrier to coordination between LTSS and medical care providers. In a case study of eight organizations financially responsible for both medical and LTSS care, only one site had a fully integrated EHR system that was accessible to both medical care and LTSS care providers. Six of the sites used separate systems for care management and medical records that are not interoperable, and one site used paper records for medical and care management services and had access to the EHR at one coordinating hospital (NCQA, 2015). This barrier to coordination was echoed by stakeholders in our interviews; they stressed the importance of the care coordinator role and the need for this person to be the communication hub between all of an individual’s providers.

McGinn-Shapiro, M., S. Mitchell, E. G. Walsh, M. Ignaczak, & L. Bercaw. (2015). Information exchange in integrated care models: final report. Available at https://aspe.hhs.gov/basic-report/information-exchange-integrated-care-models-final-report.

National Committee for Quality Assurance (NCQA). Policy Approaches to Advancing Person-Centered Outcome Measurement. 2015. The John A. Hartford Foundation and The SCAN Foundation. Available at https://www.ncqa.org/Portals/0/HEDISQM/Research/Policy%20Report_Final%20Report_TSF%202-1.pdf.

Saucier, P., and B. Burwell. (2015). Care Coordination in Managed Long-Term Services and Supports. Washington, DC: AARP Public Policy Institute. Available at http://www.aarp.org/content/dam/aarp/ppi/2015/care-coordination-in-managed-long-term-services-and-supports-report.pdf.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

Not applicable. We collected information about race and ethnicity during testing. However, due to the overall low rates, we did not believe additional analysis of disparities would provide meaningful information.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

We could not find any research on disparities in the sharing of care plans among the LTSS enrollee population. Studies have identified persistent racial and ethnic disparities regarding advanced care planning (Barwise et al., 2016; Effiong & Myrick, 2012; Garrido et al., 2014). However, most other research focuses on the identification of disparities in the need for and use of LTSS more broadly, which highlight the need for shared care plans.

The Congressional Budget Office identified racial and ethnic disparities in the need for LTSS. More specifically, it found that older black and Hispanic individuals have higher rates of functional impairment than whites (CBO, 2013).

We collected information about race and ethnicity during testing. However, due to the overall low rates, we did not believe additional analysis of disparities would provide meaningful information.

Barwise, A., M. Wilson, R. Kashyap, O. Gajic, & B. W. Pickering. (2016). Disparities in Advanced Care Planning in The ICU and End of Life Decision Making. Available at http://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2016.193.1_MeetingAbstracts.A7926.

Congressional Budget Office (CBO). (2013). Rising Demand for Long-Term Services and Supports for Elderly People. Washington, DC: Congressional Budget Office.

Effiong, A. & D. Myrick. (2012). H.R. 1589: addressing racial and ethnic disparities in advance care planning among Medicare beneficiaries. BMJ Supportive & Palliative Care, 2, 181.

Garrido, M. M., S. T. Harrington, & H. G. Prigerson. (2014). End-of-life treatment preferences: a key to reducing ethnic/racial disparities in advance care planning? Cancer, 120(24), 3981-3986.http://www.ncbi.nlm.nih.gov/pubmed/25145489

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

De.6. Non-Condition Specific*(check all the areas that apply):*

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

Measure specifications have been drafted, and are anticipated to be publicly posted on CMS's MLTSS website in early 2018. However, currently there is no link.

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

No data dictionary **Attachment:**

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Medicaid MLTSS enrollees who have a care plan (or part of a care plan) that was transmitted to their PCP within 30 days of the care plan's development or update date.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Time Period for Data:

13 months (November 30 of the year prior to the measurement year to December 31 of the measurement year).

Numerator Details:

Medicaid MLTSS enrollees who have a care plan (or part of a care plan) that was transmitted to the PCP within 30 days of the care plan's completion or update date. Evidence of a transmitted care plan should meet the following criteria:

- Who the care plan was transmitted to.
- Date of transmittal.
- The elements of the care plan that were transmitted.

Note: If the enrollee has more than one care plan developed during the measurement year, use the most recent care plan date to assess measure.

Definitions:

Care plan: A document or electronic record which identifies enrollee needs, preferences, and risks, and contains a list of the services and supports planned to meet those needs while reducing risks.

Transmitted: Care plan may be transmitted to providers via mail, fax, secure e-mail, alert in provider portal system, mutual access to an electronic health record (EHR) with notification to the PCP, or other electronic data system.

Primary Care Practitioner (PCP): A physician, non-physician (e.g., nurse practitioner, physician assistant), or primary care practice, who offers primary care medical services. Licensed practical nurses and registered nurses are not considered PCPs.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

Medicaid MLTSS enrollees age 18 years and older who had a care plan developed or updated in the measurement year.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

A systematic sample drawn from the eligible population, which includes enrollees:

- Who have a care plan developed or updated between November 30 of the year prior to the measurement year and December 1 of the measurement year.

- Who are 18 years and older as of the first day of the measurement year.

- Who are enrolled in an MLTSS plan for at least 30 days after the care plan's development or update date. If multiple care plan updates are documented for the year, determine continuous enrollment from the latest care plan update in the year.

- Who have either of the following benefits: 1) long-term services and supports: home- and community-based or 2) long-term services and supports: institution based.

Note: The denominator for this measure may be drawn from enrollees meeting the numerator criteria for a paired measure LTSS Comprehensive Care Plan and Update.

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

Exclude enrollees in the denominator who were not enrolled in an MLTSS plan for at least 30 days after a care plan's development or update date.

Exclude enrollees who refuse to have their care plan shared.

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

Exclude enrollees in the denominator who were not enrolled in an MLTSS plan for at least 30 days after a care plan´s development or update date. These are enrollees who may have left the plan before it was shared with the PCP.

Exclude enrollees for whom there is documentation of enrollee refusal to allow care plan sharing.

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

Not Applicable, no stratification.

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

Step 1. Determine the eligible population of MLTSS enrollees with a care plan developed or updated in the measurement period.

Step 2. From the eligible population, draw a systematic sample.

Step 3. From the systematic sample, remove enrollees that have documentation of refusal to allow care plan sharing.

Step 4. From the systematic sample, remove enrollees who were not enrolled 30 days after the date of care plan was development or update.

Step 5. From enrollees remaining after Step 4, identify all enrollees with a care plan for whom all or part of the care plan was transmitted to the primary care practitioner within 30 days of the care plan's development or latest update date.

Step 6. Divide the number of enrollees from Step 5 by the number of enrollees remaining after Step 4 to calculate the rate.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

MLTSS plans should identify a systematic sample of 411 enrollees who meet the eligible population criteria. The same sample may be used to calculate three paired measures:

- LTSS Comprehensive Assessment and Update

- LTSS Comprehensive Care Plan and Update

- LTSS Shared Care Plan with Primary Care Practitioner

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

Not applicable.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Management Data, Other, Paper Medical Records

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

Other: Case Management Records. Records are reviewed to determine if care plan was shared during the required time frame.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Health Plan

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Home Care, Other

If other: Long term non-acute care, home- and community-based services, health plan case

S.22. <u>COMPOSITE Performance Measure</u> - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

2. Validity – See attached Measure Testing Submission Form

LTSS_Shared_Care_Plan_Testing_Attachment_Nov28.docx

2.1 <u>For maintenance of endorsement</u>

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

2.2 <u>For maintenance of endorsement</u>

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

2.3 <u>For maintenance of endorsement</u>

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** TBD
Measure Title: LTSS Shared Care Plan with Primary Care Practitioner
Date of Submission: <u>11/8/2017</u>

Type of Measure:

☐ **Outcome (*including PRO-PM*)**	☐ **Composite – *STOP – use composite testing form***
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.**
- **For outcome and resource use measures**, section **2b3** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b5** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **instrument-based measures** (including PRO-PMs) **and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b1. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **instrument-based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; [12]

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.

15. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR <u>ALL</u> TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. <u>If there are differences by aspect of testing,</u>(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure*. *Testing must be provided for <u>all</u> the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: **(*must be consistent with data sources entered in S.17*)**	**Measure Tested with Data From:**
☒ abstracted from paper record	☒ abstracted from paper record
☐ claims	☐ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☒ other: abstracted from case management records	☒ other: abstracted from case management records

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

N/A

1.3. What are the dates of the data used in testing? September 1, 2014 to December 31, 2015

1.4. What levels of analysis were tested? (*testing must be provided for <u>all</u> the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: **(*must be consistent with levels entered in item S.20*)**	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☒ health plan	☒ health plan
☐ other:	☐ other:

1.5. How many and which <u>measured entities</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Five MLTSS plans representing both large national plans and small local plans from five states, located in geographically diverse regions of the country were included in the testing and analysis. All five participating health plans enrolled two or more of the major LTSS sub-populations, which include frail older adults age 65 and over, people under age 65 with physical disabilities, people with developmental or intellectual disabilities, and people with serious mental illness. All five participating plans covered Medicare benefits as well as Medicaid LTSS and had at least one of the following types of contracts: Medicare-Medicaid Plan (MMP), Dual Eligible Special Needs Plan (D-SNP), or Fully Integrated Dually Eligible Special Needs Plan (FIDE SNP).

1.6. How many and which <u>patients</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Each of the five plans identified a random sample of 150 enrollees from an eligible population of MLTSS members age 18 and older as of January 1, 2015. Two groups of enrollees were included in the sample identified by each plan (i.e., "new" and "established" enrollees). "New" enrollees were members who were newly enrolled between September 1, 2014 and August 31, 2015, without any gaps in enrollment during the first 120 days during the measurement year of CY 2015.

"Established" enrollees were members who were enrolled prior to September 1, 2014 and enrolled continuously with no more than one 45-day gap throughout the measurement year. To ensure that the final sample of 150 enrollees included adequate data from both subgroups, each health plan was asked to include at least 40 "New" enrollees in the sample. Forty-six enrollees were removed from the sample for inconsistent information, and the total sample included in the testing and analysis contained 715 enrollees. Table 1 summarizes the enrollees' characteristics for the sample. Table 2 summarizes their place of residence, type of plan, whether they were new or established enrollees, and presence of chronic conditions or activities of daily living (ADL) limitations.

Table 1. Analytic Sample Demographic Information

Characteristics	Percentage of enrollees in the testing sample (n=715)
Sex	
Female	68.5
Male	31.0
Missing	0.4
Age	
Under 18	0.7
18-40	6.9
41-64	33.3
65 and older	59.0
Missing	0.1
Race	
White	37.9
Black/African American	27.0
Asian	3.5
American Indian/Alaskan Native	0.3
Multi-race	0.1
Hawaiian/Pacific Islander	0.1
Unknown	19.2
Other	11.2
Ethnicity	
Non-Hispanic	55.9
Hispanic	17.3
Unknown	22.0
Primary language	
English	66.7
Spanish	10.4
Missing	17.1
Other	5.9

Source: Mathematica analysis of MLTSS enrollees from five health plans. Information from 46 enrollees was excluded due to data quality issues.

Table 2. Analytic Sample LTSS information

Characteristic	Percentage of enrollees in the testing sample (n=715)
Place of residence	
Home or community residence	77.6
Nursing facility	14.3
Assisted living facility	1.1
Other institution	0.7
Missing	6.3
MLTSS program	
Integrated plan (MMP, D-SNP, FIDE-SNP)	68.3
Non-integrated	31.6
Missing	0.1
Type of enrollee	
New	48.6
Established	51.3
Missing	0.1
Chronic conditions present by end of measurement year	
Arthritis	42.8
Asthma	13.2
Cancer	8.0
Cardiac conditions (e.g., CAD, arrhythmia)	42.1
Dementia	17.5
Depression	34.3
Diabetes	35.4
Gastrointestinal (GI) disorders (e.g., IBD, cirrhosis)	26.0
Chronic Heart Failure	16.9
HIV	1.5
Neurological Disorders	20.7
Other Pulmonary Conditions (e.g., COPD)	24.2
Psychotic Disorder	11.9
Renal Disease	13.4
Stroke	16.1
ADL Limitations present by end of measurement year	
Walking	69.5
Toileting	57.2
Bathing	61.5
Eating	26.9
Transferring	61.0
Dressing	59.0

Source: Mathematica analysis of MLTSS enrollees from five health plans. Information from 46 enrollees was excluded due to data quality issues.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

No difference in the sample size used for testing.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

No patient-level sociodemographic (SDS) variables were analyzed. All patients in the sample were Medicaid-eligible.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☒ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Data presented here is on the final measure specifications. Additional detail on the recipients and timing of the shared care plan can be found in the Appendix: Additional Testing Data.

Reliability of Data Elements

We calculated reliability of the critical data elements used in the measure with Cohen's kappa statistic to evaluate inter-rater reliability. Four participating health plans contributed thirty paired assessments and one plan contributed twenty-four paired assessments for analysis by two independent assessors. In total, the records of 144 MLTSS health plan enrollees were used for this analysis.

When comparing observations made by two individuals, Cohen's kappa statistic, or $\hat{\kappa}$ (ranging from 0 to 1) measures the percentage of agreement between individuals. If the observed agreement is greater than or equal to chance agreement, then $\hat{\kappa} \geq 0$, with $\hat{\kappa} = 1$ signifying perfect agreement. If the observed agreement is less than or equal to chance agreement, then $\hat{\kappa} \leq 0$ (Fleiss, Levin and Paik, 2003). We calculated the $\hat{\kappa}$ statistic reflecting the amount of agreement among key data elements as:

$$\hat{\kappa} = \frac{\rho a - \rho e}{1 - \rho}$$

Where ρe is the expected percent chance agreement and ρa is the observed agreement.

Reliability of Measure Rates

To examine the reliability of measure performance rates, we evaluated split-sample reliability using the Inter-class correlation coefficient (ICC) to determine agreement. Enrollees were randomly assigned to each of the two groups, and rates were calculated separately for each group and then compared at the health plan-level.

When evaluating the reliability of rates or proportions, the inter-class correlation coefficient, $\hat{\rho}$, summarizes the estimated agreement among observations. The inter-class correlation coefficient (ICC) is the ratio of the subject variance to the total variance (the subject variance plus the error variance). Higher values of $\hat{\rho}$ indicate that the subject variance exceeds the error variance by a wide margin (Gwet, 2014). We calculated the $\hat{\rho}$ statistic reflecting the amount of agreement among measure results as:

$$\hat{\rho} = \frac{\sigma_s^2}{\sigma_s^2 - \sigma_e^2}$$

where $\sigma s2$ is the subject variance, and $\sigma e2$ is the error variance.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

The single data element in the measure met the threshold for slight reliability at ($\hat{\kappa}$=.1667)

Table 3. Reliability of key data elements

Measure (elements)	Data element	Kappa statistic	Interpretation
General	Date of Birth	0.8426	Almost Perfect
(4)	Sex	0.8788	Almost Perfect
	Place of Residence	0.4706	Moderate
	Date of First Enrollment	0.7108	Substantial
	Date of First Disenrollment	-0.5052	Less than Chance Agreement
LTSS Shared Care Plan with Primary Care Practitioner	Care Plan Shared?	0.1667	Slight

Source: Mathematica analysis of paired data from 144 MLTSS enrollees, representing five health plans.

Notes: Interpretation of Kappa statistic used the following standards: < 0 – Less than chance agreement; 0 – 0.2 Slight agreement; 0.21 – 0.39 Fair agreement; 0.4 – 0.59 Moderate agreement; 0.6 – 0.79 Substantial agreement; 0.8 – 0.99 Almost Perfect agreement; 1 Perfect agreement. (Landis and Koch, 1977)

*The remainder of LTSS Shared Care Plan with Primary Care Practitioner elements were assessed too infrequently among the 144 paired assessments (<30) to allow for IRR analysis.

Reliability of Measure Rates

ICCs for the measure rate exceed 0.9, indicating almost perfect agreement between the samples, and showing a significant association at p<0.5 or less.

Table 4. Reliability of Measure Rates

Measure	ICC statistic	Interpretation
LTSS Shared Care Plan with Primary Care Practitioner		
Rate: Shared with PCP without Consideration for Share Method	0.9668**	Almost Perfect

Source: Mathematica analysis of data from 715 MLTSS enrollees, representing five health plans.

Notes: Interpretation of ICC used the following standards: < 0 – Less than chance agreement; 0 – 0.2 Slight agreement; 0.21 – 0.39 Fair agreement; 0.4 – 0.59 Moderate agreement; 0.6 – 0.79 Substantial agreement; 0.8 – 0.99 Almost Perfect agreement; 1 Perfect agreement. (Landis and Koch, 1977)

**Significantly associated at the p<0.01 level.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

The Interclass Correlation Coefficient for the measure rate exceeds 0.9, indicating almost perfect agreement between the samples for the single data element indicating that the care plan was shared. However, the other elements in the LTSS Shared Care Plan with Primary Care Practitioner measure were assessed too infrequently among the 144 paired assessments (<30) to allow for inter-rater reliability analysis.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☐ **Performance measure score**

☒ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Empiric Validity of Performance Measure Score

Empiric validity was assessed by analyzing the convergent validity (whether the measure results correlate with other theoretically related quality measure results) of this measure. Comparison with established health plan quality measures was not possible due to the lack of publicly available standardized measures for the MLTSS population. Therefore, we analyzed correlation between this care plan sharing measure with four other measures that were tested for the MLTSS population in this project:

- LTSS Comprehensive Assessment and Update
- LTSS Comprehensive Care Plan and Update Measure
- LTSS Re-Assessment and Care Plan Update after Inpatient Discharge
- Falls: Screening, Risk-Assessment and Plan of Care to Prevent Future Falls

We examined the correlation of different versions of the five MLTSS measures with each other using the Spearman Rank Correlations. As with this assessment measure, we tested several variations of each of the other MLTSS measures. Although we analyzed validity among all versions of the five measures for ease of review we present results for only the most promising versions of the measures in Table 5.

Systematic Assessment of Face Validity

By surveying several panels of stakeholders, we were able to assess face validity, which indicates whether the measure accurately represents the concept being measured and achieves the purpose for which it is intended (i.e., to measure the percent of beneficiaries being assessed and the quality of the assessments conducted).

In November 2017, 13 members of the 2017 Technical Expert Panel (see Ad.1 for member list) voted on the face validity of the measure via a web survey. The TEP voted on whether they “strongly agree”, “agree”, “disagree”, or “strongly disagree” with the following survey items:

1. Denominator is appropriate given the intent of the measure

2. Numerator Rate is appropriate given the intent of the measure

3. Exclusion is appropriate given the intent of the measure

4. Would high performance on this measure indicate that a health plan is providing higher quality care?

5. In the future, do you think that performance scores on this measure will distinguish between good and poor performance?

TEP members had the opportunity to provide written comments on all the survey items above, as well as an additional open response question:

6. Do you have any recommendations that would help strengthen the Shared Care Plan with Primary Care Practitioner Measure?

See Appendix 1 for summaries of the written comments for each survey item. While the majority of TEP members agreed with the measure, the comments included in the summaries are largely from those who disagreed or suggested revisions to the measure.

Additional Face Validity Feedback

Under a previous CMS contract (Contract No. HHSM-500-2010-00026I/HHSM-500-T0011), a multistakeholder technical expert panel (TEP) of 20 individuals representing home and community based services organizations, disability service organizations, organizations providing services for aging populations, state Medicaid and quality organizations, and consumer advocacy groups was convened to provide input on the MLTSS measure development and testing processes (see Ad.1 for TEP member list – 2013 TEP). Under the current contract, we convened a new TEP with a similar size and composition to inform the continued development and testing of this measure (see Ad.1 for TEP member list – 2017 TEP). Six members of the 2017 TEP also participated in a workgroup that provided in-depth feedback on the MLTSS measures specifically (2017 MLTSS Workgroup).

We also received feedback from a three-week public comment period hosted on CMS's online public comment system. The public comment period was open and broadcast to all interested parties.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Empiric Validity of Performance Measure Score

Among the recommended measure rates, very few positive, significant relationships were observed. The Spearman Rank Correlation coefficient, $\hat{\rho}$, showed no significant relationship between the LTSS Shared Care Plan with Primary Care Practitioner measure and any other MLTSS measure as shown in Table 5 (correlation of recommended measure rates).

Table 5. Correlation of recommended measure rates

Measures	MLTSS-1, Rate 3	MLTSS-1, Rate 4	MLTSS-2, Rate 3	MLTSS-2, Rate 4	MLTSS-3	MLTSS-4, Rate 1	MLTSS-4, Rate 2	MLTSS-5 Screening	MLTSS-5 Assess	MLTSS-5 Care Plan
MLTSS-3, Rate: Shared with PCP, w/o consideration for Share Method	-0.0574	-0.0574	0.5000	0.5000	--	-0.4472	-0.4472	-0.8944*	-0.6250	NA

Source: Mathematica analysis of data from 715 MLTSS enrollees, representing five health plans.

*Significant association, at $p < 0.05$

MLTSS-1 = LTSS Comprehensive Assessment and Update

Rate 3 = Assessment with 9 core elements documented

Rate 4 = Assessment with at least 12 supplemental elements documented

MLTSS-2 = LTSS Comprehensive Care Plan and Update

Rate 3 = Care plan with 7 core elements documented

Rate 4 = Care plan with at least 4 supplemental elements documented

MLTSS-3 = LTSS Shared Care Plan with Primary Care Practitioner

Rate = Care plan transmitted to provider within 30 days of the plan's development or update

MLTSS-4 = LTSS Re-Assessment and Care Plan Update after Inpatient Discharge

Rate 1 = Re-assessment and care plan update, no face to face requirement

Rate 2 = Re-assessment only, no face to face requirement

MLTSS-5 = Falls: Screening, Risk-Assessment, and Plan of Care to Prevent Future Falls

Rate 1 = Screened for falls risk

Rate 2 = Falls risk assessment

Rate 3 = Falls risk plan of care

Systematic Assessment of Face Validity

Table 6 contains the voting results from the survey. Overall, the majority of TEP members supported the denominator, numerators, and exclusions for the Long Term Services and Supports (LTSS) Shared Care Plan with Primary Care Practitioner measure. The majority of TEP members also agreed that the measure is reflective of quality and had the potential to distinguish performance.

Table 6

Response	Denominator is appropriate given the intent of the measure	Numerator Rate is appropriate given the intent of the measure	Exclusion 1 is appropriate given the intent of the measure	Would high performance on this measure indicate that a health plan is providing higher quality care?	In the future, do you think that performance scores on this measure will distinguish between good and poor performance?
Strongly Agree	1	1	2	1	1
Agree	11	10	10	7	6
Disagree	0	0	1	2	3
Strongly Disagree	0	2	0	3	3
No response	1	0	0	0	0
Total % Agree	92%	85%	92%	62%	54%

Additional Face Validity Feedback

The TEP noted that his measure is a first step and is not meant to be used alone to determine quality; rather, it is one piece of critical care delivery in determining whether individuals needs and preferences are being assessed and addressed in care plans - including non-medical needs. The measurement team agrees that sharing a care plan with a PCP is an important initial activity to determine that organizations are assessing the needs of beneficiaries, and as documentation of the core elements improves among plans over time, performance is likely to become more meaningful and useful.

Additionally, feedback from the public comment period was generally supportive of the measure. Several comments stressed the importance of harmonizing the measures with existing measures and tools. Numerous comments offered ways that the measure specifications could be furthered clarified, including denominator and numerator criteria, timing, construction of required elements, setting of the assessment and exclusion criteria. A few comments asked us to explore the importance of the measure and the burden for health plans to report during testing.

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Empiric Validity of Performance Measure Score

As documentation of the core elements improves among plans over time, performance is likely to become more meaningful and useful and the internal validity of the measures should also improve accordingly.

This measure assesses the percentage of enrollees age 18 and older whose care plan was shared with their PCP within 30 days of development or update. This measure does not require documentation of the care plan's transmission method. Testing results indicate marked variation in care plan sharing practices. The refined measure specification that focuses on sharing of information with the PCP should represent a reasonable benchmark that plans can work to meet

over time. As performance improves, it may be useful to add back the requirement for documentation of the method for sharing the care plan; however, we do not recommend doing so at this time.

Testing results primarily highlighted the overall sub-optimal performance even for the simplified measure (all but two plans had rates of zero). However, testing results suggested the existence of meaningful variation (one plan with meaningful and significantly higher performance). Participants in the TEP, the workgroup and public commenters all supported moving forward with this measure.

Systematic Assessment of Face Validity

The voting results suggest that this is a valid measure. TEP members who did not support the measure cited that this measure is a process measures, and an outcome measure would better serve this population. The measurement team agrees that outcome measures are a long-term goal, but before useful data on outcomes can be collected, organizations must assess and document the needs of beneficiaries using a standard set of guidelines. One cannot measure an outcome that is not being assessed in the first place. This measure is a first step and is not meant to be used alone to determine quality; rather, it is one piece of critical care delivery in determining whether individuals' needs and preferences are being assessed and addressed in care plans - including non-medical needs.

Additional Face Validity Feedback

Stakeholder input suggests that this measure is valid for assessing the need for long term services and supports for plan enrollees. The findings from the TEPs, workgroup, and public comment suggest that modifications made to the measure specification produced valid results.

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section* 2b3

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

N/A

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

N/A

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

***If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section* 2b4.**

2b3.1. What method of controlling for differences in case mix is used?

☒ No risk adjustment or stratification

☐ Statistical risk model with _risk factors

☐ Stratification by _risk categories

☐ Other,

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

N/A

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

N/A

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

N/A

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?

N/A

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

N/A

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

N/A

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

N/A

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

N/A

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

N/A

2b3.9. Results of Risk Stratification Analysis:

N/A

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i*.e., what do the results mean and what are the norms for the test conducted*)

N/A

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

N/A

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

To evaluate whether the measures demonstrated statistically significant variation in performance across health plans and/or showed sub-optimal performance (suggesting room for improvement), we calculated the sample mean, minimum and maximum values, and standard deviations for the five health plans-level results. We utilized two-tailed T-tests to evaluate whether each health plan's results differed significantly from the sample mean.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

MLTSS plan performance on the measure is presented in Table 6. Health plan 01 demonstrated rates that differed significantly from the mean at the .05 level.

Table 6. Performance rates by health plans with significant difference noted

Health Plan	Care Plan Shared with PCP
HP 01	23.4*
HP 02	0.0
HP 03	9.2
HP 04	0.0
HP 05	0.0
Minimum	0.0
Mean	6.5
Maximum	23.4
Standard deviation	10.2

Source: Mathematica analysis of enrollee data from 5 MLTSS plans.

Note: Health plan identifiers "HP 01" through "HP-05" are used to protect the confidentiality of health plans participating in beta testing.

*Significantly different from the mean at the .05 level, two-tailed test

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

Although we observed poor overall performance, health plan 01 had performance rate that demonstrated a statistically significant difference from the mean. This finding indicates that the measure as specified does have the potential to identify meaningful differences in performance between health plans.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

***If only one set of specifications, this section can be skipped*.**

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the*

measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator). ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable. There is only one set of specifications for this measure.

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

N/A

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

N/A

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

The primary focus of this measure is an assessment of missing data – i.e., whether MLTSS enrollees show documentation of information relating to sharing of a care plan with the primary care provider. When required elements are missing, the enrollee is considered not to have met the numerator requirements. Therefore, the measure results themselves reflect the extent and distribution of missing data overall and by plan. The extent of missing data for key data elements is described in further detail in the Additional Testing Results Appendix.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; <u>if no empirical sensitivity analysis</u>, identify the approaches for handling missing data that were considered and pros and cons of each*)

Please see details in the Additional Testing Results Appendix.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; <u>if no empirical analysis</u>, provide rationale for the selected approach for missing data*)

The low rates for this measure reflect the lack of standardization regarding practices for sharing care plans with primary care providers. This measure assesses the percentage of Medicaid MLTSS enrollees who have care plan shared with a primary care provider within 30 days of development or update, and in doing so, should help address this lack of standardization.

Appendix: Additional Testing Data

<u>Methods:</u>

Prior to testing reliability and validity, we assessed the individual elements included in the measure and tested alternative versions of calculating the measure.

1) We analyzed how often new or updated care plans were shared with key LTSS providers and PCPs, and the number of days between when the care plan was first developed or updated and then shared.
2) We analyzed which types of providers were more likely to be sent care plans. The measure as specified prior to beta testing required documentation of sharing the care plan with both the PCP and at least one key LTSS provider.
3) Finally we analyzed who the care plan with was shared by exploring three different possible rates.
    - Original Rate: Percentage of enrollees with a care plan transmitted to the primary care practitioner (PCP) and other key LTSS providers
    - Version 1 - PCP Only: Percentage of enrollees with a care plan transmitted to the primary care practitioner (PCP) with documentation of the method of the care plan's transmission (e.g., fax, email, EHR)
    - Version 2 - PCP only - no requirement of transmission method: Percentage of enrollees with a care plan transmitted to the primary care practitioner (PCP) without any documentation of method of transmission requirement

Results:

Frequency of Shared Care Plan

Approximately 30 percent of the enrollees in the measure's denominator (133 enrollees) had their care plan shared with a PCP or a key LTSS provider at least once in the measurement period (Table 1). Among all the 133 enrollees who had an initial care plan or care plan update shared, 63 percent were shared within 30 days, and most were shared within 10 days.

Table 1. Care plan sharing rates by enrollee type

Enrollees with	Frequency	Percentage of enrollees with a care plan eligible for MLTSS-3 (n =438)
Documentation of a care plan shared with an eligible provider (PCP or Key LTSS)	**133**	**30.4**
Care plan shared within 30 days from its creation	84	19.2
Care plan shared after 30 days of its creation	21	4.8
Data entry error*	28	6.4
No documentation of a care plan shared with an eligible provider	**305**	**69.6**

Source: Mathematica analysis of enrollee data from 5 MLTSS plans.

Types of Providers Care Plan was Shared With

We also analyzed which types of providers were more likely to be sent care plans. The measure as specified prior to beta testing required documentation of sharing the care plan with both the PCP and at least one key LTSS provider.

Of the 133 enrollees that had an initial care plan or care plan update shared, 58 percent of the enrollees had the care plan shared with the enrollee's PCP (Table 2), while only 9 percent of these enrollees had the care plan shared with both the PCP *and* one or more LTSS providers. This represents 18 and 3 percent, respectively, of all 438 enrollees in the measure's denominator. Among enrollees with share care plans or care plan updates, "other provider" was the most frequently selected choice (35 percent) for key LTSS providers receiving the care plan, followed by providers of personal care in the home (14 percent).

Regarding the care plan's transmission method, among enrollees with an initial care plan or care plan update shared, approximately 50 percent had the method of sharing (fax, email, EHR) documented in the enrollee's record. This represents only 15 percent of all enrollees in the measure's denominator. The most common method for transmitting a

care plan or care plan update was through fax, while the least common was a notification through the electronic health record system.

Table 2. Documented type of provider with whom the care plan or care plan update was shared

Type of provider receiving the shared care plan	Frequency	Percentage of enrollees with a shared care plan (n =133)
Care plan or care plan update shared with Primary Care Practitioner (PCP)	**77**	**57.9**
Care plan or care plan update shared with LTSS Provider		
Other Provider*	47	35.3
Provider of Personal Care in the Home	19	14.2
Provider of Residential and Habilitation Center	3	2.3
Physical or Occupational Therapist	0	0.0
Skilled Nurse	0	0.0
Care plan or care plan update shared with PCP and one LTSS Provider	**12**	**9.0**

*Primarily LTSS provider agencies and adult day programs

Source: Mathematica analysis of enrollee data from 5 MLTSS plans.

Health Plan Performance

While our results indicate that there is documentation of components for the majority of care plans that are shared (85 percent), the percentage of care plans actually shared with the PCP was quite low, with only one health plan showing a non-zero measure rate (Health Plan 3) at about 3 percent for the LTSS Shared Care Plan with Primary Care Practitioner Original Rate (Table 2) during beta testing.

Table 2. Performance rates by health plan

Health Plan	Original Rate: Shared with PCP and key LTSS Providers	Version 1: Shared with PCP only, required documentation of method of transmission	Version 2: PCP only - no requirement of transmission method
HP 01	0.0	0.0	23.4
HP 02	0.0	0.0	0.0
HP 03	3.1	7.7	9.2
HP 04	0.0	0.0	0.0
HP 05	0.0	0.0	0.0
Minimum	0.0	0.0	0.0
Mean	0.6	1.5	6.5
Maximum	3.1	7.7	23.4
Standard deviation	1.4	3.4	10.2

Source: Mathematica analysis of enrollee data from 5 MLTSS plans.

Note: Health plan identifiers “HP 01” through “HP-05” are used to protect the confidentiality of health plans participating in beta testing.

MLTSS-3 = LTSS Shared Care Plan with Primary Care Practitioner

After discussing these results with advisory panel stakeholders, we determined that plan abstractors found the term “key LTSS provider” subjective and confusing. We then surveyed workgroup members to determine which types of providers should always be sent care plans for MLTSS enrollees. Among the six workgroup members who completed the

survey, PCPs ranked highest as the provider type that should always receive the care plan (67 percent), followed by personal care assistants (50 percent), and skilled nurses (33 percent). We also noted, verification of which key LTSS providers enrollees were currently receiving care (and therefore should have their care plan shared with) was not possible in the current measure (i.e., there was no way to identify if the enrollee had a skilled nurse who should be receiving the care plan).

Interpretation:

These results, in combination with challenges associated with defining "key LTSS providers," led to a recommendation to limit the measure to focus only on sharing the care plan with the PCP, without any requirement as to which components or elements must be shared (Version 2 above).

Systematic Assessment of Face Validity: Summaries of Written Feedback

1. Denominator is appropriate given the intent of the measure

The majority of TEP members (12 out of 13, or 92%) agreed that the denominator is appropriate for the measure. One TEP member asked for clarification as to how MLTSS plan enrollees were defined, as LTSS state plan benefits can differ from HCBS waiver services.

2. Numerator Rate is appropriate given the intent of the measure

The majority of TEP members (11 out of 13, or 85%) agreed that the numerator is appropriate for the measure. Comments from the TEP included the suggestion that a summary of the care plan would be more efficient for the PCP, as care plans can be lengthy. While one TEP member noted that for LTSS measures the PCP may not be involved in the execution of the care plan, the measurement team agrees that the PCP should be informed regarding the care plan.

3. Exclusion 1 is appropriate given the intent of the measure

Most TEP members (12 out of 13 or 92%) thought that the preferences of enrollees regarding the sharing of a care plan with the PCP was an appropriate exclusion for this measure. One TEP member noted that this exclusion was consistent with recognizing the rights of enrollees to make decisions regarding their own healthcare.

4. Would high performance on this measure indicate that a health plan is providing higher quality care?

While most of the TEP (8 out of 13, or 62%) responded that good performance on this measure is indicative of high quality care, several TEP members had comments regarding the disposition of care plans once transmitted to the provider. Members noted that sharing a care plan with a PCP was not an indicator that the PCP had reviewed the care plan, or the plan was sent to the appropriate provider, or that members were receiving needed services specified in the plan. While we agree with the members that this measure does not indicate whether the PCP acted on the care plan, sending the care plan is first step of coordination between LTSS and medical care providers and is critical to ensuring PCPs are aware of the care being provided by the LTSS providers. Unfortunately, it is not possible to determine with current documentation whether the care plan was reviewed by the PCP.

Another TEP member questioned if all LTSS services outlined in a care plan are healthcare related and appropriate to share with the provider. To address this concern the measure was specified to allow only a portion of the care plan to be transmitted to the PCP, which will allow for the MLTSS plan to determine which components of the care plan are most relevant for the PCP.

5. In the future, do you think that performance scores on this measure will distinguish between good and poor performance?

The TEP was mostly (7 out of 13, or 54%) supportive of the measure's ability to distinguish between good and poor performance. Commenters noted that the measure was specified as a process measure and does not correlate patient outcomes to a care plan. While the measurement team agrees with the value of outcome measures for this population, MLTSS is an evolving area of measurement and process measures can drive the standardization of data needed to develop outcome measures. TEP members that supported the measure were in agreement that sharing a care plan with the PCP is a first step in improving performance.

6. Do you have any recommendations that would help strengthen the Long Term Services and Supports (LTSS) Shared Care Plan with Primary Care Practitioner measure?

TEP members that were supportive of the measure noted that this was a good first step to ensure that plans were reaching out to the PCP. One TEP member suggested extending the 30-day transmission window to 45 days. One TEP member noted that some percentage of enrollees in managed LTSS plans do not have a primary care physician. One TEP member suggested that there was greater importance for the patient or consumer to receive a copy of the care plan than the PCP.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

Some data elements are in defined fields in electronic sources

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

This measure is primarily collected from health plan and case management records, many of which are electronic. However, the data elements needed for this measure are not currently standardized (i.e., data elements may be in free text instead of structured fields). As the LTSS field moves forward, we anticipate the data elements needed to calculate this measure will become structured allowing for an eMeasure in the future.

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

Not applicable.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

Not applicable, no fees or licensing are currently required.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
Public Reporting Quality Improvement (Internal to the specific organization)	

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

Not applicable.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

Not applicable; the measure is under initial endorsement review and is not currently used in an accountability program. A measure implementation plan will be developed by, or in conjunction with, CMS.

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

This measure is intended for use by states and health plans to monitor and improve the quality of care provided for the Medicaid MLTSS enrollee population. States and health plans may choose to begin implementing the measures based on their programmatic needs.

In May 2017, the National MLTSS Health Plan Association recommended a set of model LTSS performance measures and network adequacy standards in an effort to assist states in complying with the 2016 final rule on managed care in Medicaid and Children's Health Insurance (CHIP). The Long Term Services and Supports (LTSS) Shared Care Plan with Primary Care Practitioner measure is included in the set of recommended measures that assesses person-centered planning and coordination.

http://mltss.org/wp-content/uploads/2017/05/MLTSS-Association-Quality-Framework-Domains-and-Measures-042117.pdf

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

Not applicable.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

Not applicable.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

Not applicable.

4a2.2.2. Summarize the feedback obtained from those being measured.

Not applicable.

4a2.2.3. Summarize the feedback obtained from other users

Not applicable.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Not applicable.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

This measure is being considered for initial endorsement. Adoption of this performance measure has the potential to improve the quality of care for Medicaid enrollees who are receiving long-term services and supports.

Delivering effective care coordination for complex populations, such as Medicaid MLTSS enrollees, begins with conducting and regularly updating a comprehensive assessment to identify enrollees' needs, developing and regularly updating an individualized care plan to indicate the specific services and supports to be provided, and sharing care plans to inform care team members of services that should be coordinated. State Medicaid agencies have implemented numerous Medicaid MLTSS care coordination models. Most models require the development of a care plan at initial enrollment and on a regular basis thereafter, as well as the use of team based care to implement the care plan. Similarly, numerous other programs that deliver care to individuals who are similar to (or in some cases the same as) Medicaid

MLTSS enrollees require care plans and team based care, including patient-centered medical homes, Medicare managed care plans (e.g., Special Needs Plans, Financial Alignment Initiative dual eligible enrollee demonstration plans), state Medicaid home and community-based services 1915(c) waiver programs, the Program for All-Inclusive Care for the Elderly (PACE), and Medicaid Health Homes. In order for team-based care to be effective, providers must share the care plan and communicate changes and updates to the care plan so that all members of the care team have a complete picture of the person's needs, preferences, and services and supports provided.

This measure would address the existing gap in information sharing by assessing the percentage of Medicaid LTSS enrollees who have a care plan developed and shared with their PCP. A standardized measure of care plan sharing will allow for apples-to-apples comparisons of LTSS plans across states.

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

Not applicable. This measure has not been implemented yet. There were no unexpected findings identified during testing of this measure.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

Not applicable. This measure has not been implemented yet. There were no unexpected findings identified during testing of this measure.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

No

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

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?

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

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);

OR

Multiple measures are justified.

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

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment Attachment: LTSS_Shared_Care_Plan_Additional_Testing_Results_Nov28.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare & Medicaid Services, Centers for Medicaid & CHIP Services

Co.2 Point of Contact: Roxanne, Dupert-Frank, Roxanne.Dupert-Frank@cms.hhs.gov, 410-786-9667-

Co.3 Measure Developer if different from Measure Steward: Mathematica Policy Research

Co.4 Point of Contact: Henry, Ireys, hireys@mathematica-mpr.com, 202-554-7536-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

2017 Technical Expert Panel

Carol Raphael, Manatt Health Solutions (Chair)

Ann Hwang, MD, Community Catalyst

Ari Houser, PhD, AARP Public Policy Institute

Dennis Heaphy, MPH, Disability Policy Consortium

Joe Caldwell, PhD, National Council on Aging

Lauren Murray, BA, National Partnership for Women and Families

Maggie Nygren, EdD, American Association for People with Disabilities

RoAnne Chaney, MPA, Michigan Disability Rights Coalition

Mary Lou Bourne, National Association of State Directors for Developmental Disabilities Services

Raina Josberger, MS, New York State Department for Health

Jason Rachel, PhD, Virginia Department of Medical Assistance Services

Balu Gadhe, MD, CareMore

Patricia Kirkpatrick, MJ, RN, CPHQ, Amerigroup Corporation

Cheryl Phillips, MD, LeadingAge

Diane McComb, MSEd, American Network of Community Options and Resources

Steve Guenthner, BS, Almost Family, Inc.

Bonnie Marsh, RN, BSN, MA, Health Services Advisory Group

Brian Abery, PhD, University of Minnesota

Lisa Iezzoni, MD, Harvard Medical School

Pamela Parker, MPA, Independent Consultant-Integrated Care

Valerie Bradley, MA, Human Services Research Institute

Quality Measure Development (QMD) – Medicare-Medicaid (Dual) and Medicaid-Only Enrollees MLTSS Workgroup, January 2017

Laura Brannigan, GuildNet

Jennifer Clark, Centene Corporation

Camille Dobson, NASUAD

Patricia Kirkpatrick, Amerigroup

Michael Monson, Centene Corporation

Lauren Murray, National Partnership for Women and Families

Pamela Parker, Independent Consultant-Integrated Care

Carol Raphael, Manatt Health Solutions

2013 Technical Expert Panel

Anne Cohen, Health and Disability Policy Consultant, Disability Health Access, LLC

Patti Killingsworth, Assistant Commissioner and Chief of LTSS, Bureau of TennCare

Jennifer Lenz, Executive Director, State and Corporate Services, Health Services Advisory Group

Bonnie Marsh, Executive Director, State and Corporate Services, Health Services Advisory Group

Diane McComb, ANCOR Liaison with State Associations

Margaret A. Nygren, Executive Director and CEO, American Association on Intellectual and Developmental Disabilities

Joseph Ouslander, Professor of Clinical Biomedical Science, Florida Atlantic University

Pamela J. Parker, Manager, Special Needs Purchasing, State of Minnesota Department of Human Services

Cheryl Phillips, Senior VP Public Policy and Advocacy, Leading Age

D.E.B. Potter, Senior Survey Statistician, Agency for Healthcare Research and Quality

Juliana Preston, Utah Executive Director, HealthInsight

Genie Pritchett, Sr. Vice President Medical Services, Colorado Access

Alice Lind, Aging and Long Term Support Division, Washington State Department of Social and Health Services

The 2017 TEP and Workgroup reviewed feedback obtained during public comment, as well as alpha and beta testing results, and advised on the refinements of the technical specifications. The 2013 TEP advised on the development of the initial measure concept and preliminary specifications.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released:

Ad.3 Month and Year of most recent revision:

Ad.4 What is your frequency for review/update of this measure? Not applicable.

Ad.5 When is the next scheduled review/update for this measure?

Ad.6 Copyright statement:

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: Please include Jessica Ross (jross@mathematica-mpr.com) on any communications about these measures, as well as Roxanne Dupert-Frank and Henry Ireys. Thank you.