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 #: 3453

Measure Title: Continuity of care after inpatient or residential treatment for substance use disorder (SUD)

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

Brief Description of Measure: Percentage of discharges from an inpatient or residential treatment for substance use disorder (SUD) for Medicaid beneficiaries, ages 18 to 64, which was followed by a treatment service for SUD. SUD treatment includes having an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth encounter, or filling a prescription or being administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

Developer Rationale: Continuity of care helps to sustain the gains attained in initial treatment and to prevent relapses. There is general agreement that continuity of care (including encounters with the health system within a defined period of time) after discharge from inpatient or residential care for SUD is related to better outcomes including reduced substance use (DeMarce, Lash, Stephens, Grambow, & Burden, 2008; McKay & Hiller-Sturmhofel, 2011), readmissions (Mark et al., 2013; Reif et al., 2017), and criminal justice involvement (McKay, 2009), lower risk of death in the two post-discharge years (Harris et al., 2015), and improved employment status (McKay, 2009).

Although the definition of continuity varies across studies (including outpatient treatment, inpatient or residential treatment, pharmacotherapy), the findings are consistent. In spite of the benefits of continuity of care, many patients do not receive treatment services after being discharged from residential or inpatient hospital care (Garnick, Lee, Horgan, Acevedo, & Washington Circle Public Sector, 2009; Harris et al., 2006; Rubinsky et al., 2017). Continuity after inpatient or residential treatment has been found to be generally low and the variation in continuity rates suggests that there is substantial opportunity for improvement.

- A study using claims data for Medicaid enrollees ages 18-64 who had an inpatient hospital or residential detoxification admission for SUD found that 75% of these enrollees did not receive at

least one continuity of care services (outpatient, intensive outpatient, or residential) within 14 days of discharge. This varies by state and ranged from 50% to 83% (Reif, Acevedo, Garnick, & Fullerton, 2017).

- In five states' public sector SUD treatment systems, continuity of care rates for receiving at least one residential or outpatient treatment service within 14 days after discharge, ranged from 15% to 60% (Garnick et al., 2009).
- In a study with veterans, continuity of care was defined as engaging in continuing care, i.e., the total number of consecutive months after intensive treatment (e.g., inpatient or residential) during which the patient had 2 or more clinic visits and no inpatient SUD or psychiatric readmissions. Only 32% of patients had two or more continuity of care visits during the month after discharge. At three months after discharge from inpatient/residential only 17% of patients had continuity care and this was further reduced to 10% at six months after discharge (Schaefer, Ingudomnukul, Harris, & Cronkite, 2005).
- Another study found continuity of care after residential treatment to range from 4% to 91% for SUD care within 30 days of discharge, among Veterans Health Administration (VHA) programs (Rubinsky et al., 2017). Continuity of care was defined in this study as having at least one outpatient SUD or mental health visit within 30 days. SUD and mental health "visits" were defined as clinical encounters within a SUD or mental health clinic, respectively, and included individual and group encounters as well as teleheath care.
- Yet another study of VHA residential treatment programs found that more than half (59%) of patients did not engage in two or more outpatient SUD treatment visits during the first month after discharge (Harris et al., 2006).

Use of a performance measure to identify patients who are less likely to have continuity of care or have shorter stays in follow-up care is critical in targeting extra efforts in engaging these patients (Harris, McKellar, Moos, Schaefer, & Cronkite, 2006).

Numerator Statement: Discharges in the denominator with an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth or filled a prescription for or were administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

Denominator Statement: Adult Medicaid beneficiary discharges from inpatient or residential treatment for SUD with a principal diagnosis of SUD during from January 1 to December 15 of the measurement year.

Denominator Exclusions: Exclude from the denominator for both rates:

- Discharges with hospice services during the measurement year
- Both the initial discharge and the admission/direct transfer discharge if the admission/direct transfer discharge occurs after December 15 of the measurement year.

Discharges followed by admission or direct transfer to inpatient or SUD residential treatment setting within 7 or 14-day continuity of care period. These discharges are excluded from the measure because transfer, hospitalization or admission to residential treatment within 7 or 14 days may prevent a continuity of care visit from taking place. An exception is admission to residential treatment following discharge from inpatient treatment; we do not exclude these admissions, because continuity into residential treatment after inpatient treatment is considered appropriate treatment.

Measure Type: Process

Data Source: Claims

Level of Analysis: Population : Regional and State

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

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

- The logic model shows credible downstream benefits of continuity of care including reduction in hospital admissions, substance use, and criminal justice activity, as well as improved employment status. Continuity of care also is credibly connected to sustaining treatment beyond the initial and early visits.
    - The developer noted potential intermediate financial impacts (costs of coordinating new referrals and costs to Medicaid if clients receive more services) which are expected to be offset by long-term savings.
- The evidence on which this measure is based is the result of a targeted review of the scientific literature on the benefits of continuity of care (including any encounter with the health system within a defined period of time) after discharge from inpatient or residential care for SUD and the extent to which this occurs.
- The developer cites three review articles that reported that receiving more types of services, longer duration, or more frequent visits during continuity of care treatment is related to improved outcomes.
- The table, Studies of Continuity Impact on Mortality, Substance Use, Retention in Treatment and Readmission, summarizes evidence from five studies (including 2 RCTs) related to improved outcomes in these areas.
    - Harris et al. 2015 found a 23% reduction in the odds of mortality over two years for patients with continuity of care services. Reif et al. (2017) found a 39-50% reduction in

the risk of readmission for those who received residential treatment and medication-assisted treatment. Mark et al. (2013) revealed that for every 1% increase in patients receiving follow-up care at a CMHC, there was a 5% reduction in the probability of readmission.

  - Two randomized studies supported the importance of continuity of care for those with SUDs (DeMarce et al., 2008; Garner et al., 2010). Near term (within 14 days) follow-up effects noted, absent effect sizes, were longer duration care continuity and better recovery status (OR=1.92, p<.05).

- The developer noted that greater performance on the measure has the potential to reduce health care costs, but noted that estimates should be viewed with caution.
  - The saving analyses and calculation were based on Marks et al.'s 2013 data. The percentage point savings calculated of $7.6 million should partially be offset by outpatient expenditures.

Questions for the Committee:

- Studies vary in the definitions of continuity and outcomes. Are the measure specifications reasonable based on the evidence cited (e.g., does evidence support using both the 7-day and 14-day follow-up period as the best way to measure continuity of care; is SUD treatment appropriately specified?)
- Is the evidence directly applicable to the process of care being measured (i.e., initial continuity in the first two weeks after discharge)?

Guidance from the Evidence Algorithm

Process measure based on empirical evidence but without systematic review/grading of the evidence (Box 7) → Empirical evidence includes all studies in body of evidence (Box 8) → Submitted evidence indicates high certainty that benefits outweigh undesirable effects (Box 9) → Moderate

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

RATIONALE:

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

Maintenance measures – increased emphasis on gap and variation

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

- Studies found that 14-day continuity of care rates ranged from 15 to 83%, with one study finding that an average of 25% of Medicaid enrollees received at least one continuity of care service. Other studies defined continuity of care differently, but also found a relatively low number of individuals engaged in follow-up care.
- The developer used 2014 Medicaid data to test the measure in 13 states. The overall 7-day continuity of care rate, across all states tested, was 18.4% and ranged from 8.9% in State D to 41.0% in State L. The overall 14-day continuity of care rate was slightly higher at 24.2% since the

window of time to have a continuity of care service was extended. The range for the 14-day rate was 13.2% in State D to 51.1% in State L.

- Continuity of care after inpatient or residential treatment is generally low, and the variation in continuity rates jointly suggest that there is room for improvement.

Disparities

- The developer tested the 7- and 14-day measure for the overall population in states included in testing and by Medicaid beneficiary category, age, gender, race/ethnicity, and rural/urban.
- Across all states, there were higher continuity rates for one small Medicaid eligibility subgroup described as "Medicaid-aged" (not otherwise defined), females, White beneficiaries and beneficiaries of Other race/ethnicity groups, and those from rural areas, although there were some differences across states.

Questions for the Committee:

- Does the gap in care provided warrant continuation of this national performance measure?
- Is the data provided from 2014 current enough?
- Is the developer's approach to identifying and addressing disparities appropriate?

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:

** Yes, there is evidence to support this measure. Follow up after discharge from SUD treatment is extemely poor.

** Strong theoretical and empirical evidence to support measure.

** Evidence is limited, there is not sudies that directly looked at unintented consquences. However, limited fairly direct.

** Meets evidence. No New studies I am aware of.

** This is a process measurement, and is applied directly. F/U after SUD treatment is a measure of continuity of care.

1b. Performance Gap
Comments:

** Yes, performance data on the measure is provided. Good that they are measuring at both 7 and 14 days. Though it would have been better to include measurement at 30 days post D/C too.

** Yes, there is significant variation by state.

** There is a gap in care. but again limited evidence was presented.

** Yes there is a performance gap.

** Performance gap data was presented, it included various population subgroups.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability Missing Data

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

Review A

Evaluation of Reliability and Validity:

- Medicaid Analytic Extract Data from 2014 was used that included 13 unidentified states. Medicaid enrollment numbers ranging from 100,000 to 6.2 million indicated that the population range is representative of states with broad populations.
- Apparent use of inpatient SUD services was well below 1% of the 8 million adult (age 18-64) enrollees studied. The final analytic sample is over 58,000 for both the 7- and 14-day measures, and is more fully described in Table 2 of the testing document.

Reliability

- Reliability testing was analyzed at the score level using signal-to-noise ratio.
- R values were >=0.9 across states for both the 7- and 14-day rates.

Validity

- Performance score validity testing was performed by comparing the #3453 with two measures of similar concepts: #3312 and #0576.
    - Results indicated moderate, positive, but not significant correlations between the measure and the two external measures. There was large variability in the correlations assessed.
- The TEP determined the measure has face validity, but two TEP members only supported the 7-day measure.
- Additional analyses showed variation in continuity of care rates between subgroups: age, sex, race/ethnicity, Medicaid beneficiary categories, and rural/urban.

Questions for the Committee regarding reliability:

- Is the Committee satisfied with the application and description of the between versus within state variability?
- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., analysis used, exclusions)?
- Is the absence of significant convergent validity (with other similar measures) a concern?

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

Evaluation A: Scientific Acceptability

Measure Number: 3453

Measure Title: Continuity of care after inpatient or residential treatment for substance use disorder (SUD)

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☒ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☐ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☒ **Other:** State

Measure is:

☒ **New** ☐ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**
   - It would be useful for the developers to describe how they obtained the value sets. Can the developers better support that the value sets are sensitive and specific to the concepts at hand? Reviewing the codes included indicated they have reasonable face validity, but how was that assured (e.g., TEP review, previous measure lists from NCQA, publications)?

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**

5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**
   - State-level testing was done (13 states included) using Medicaid Analytic eXtract (MAX) Medicaid files.
     - The final sample for 7-day continuity is 59,821 beneficiaries and for 14-day continuity is 58,900 beneficiaries.
   - Signal-to-noise reliability testing method was used.
     - The signal-to-noise ratio (SNR) statistic, R (ranging from 0 to 1), summarizes the proportion of the variation between state scores that is presumed to be due to real differences in underlying quality of care as opposed to background-level or random variation.

   Submission document: Testing attachment, section 2a2.2

7. **Assess the results of reliability testing**
   - The Adams' R Score for the 7-day measure across all 13 states was 0.989 and ranged from 0.970 to 0.999. The R score for the 14-day measure averaged 0.990 and ranged from 0.975 to 0.999.
   - R scores suggest that nearly all of the variability for both the 7- and 14-day measure rates was evident between rather than within states.
   - All states achieved a SNR reliability estimate >= 0.90, this suggests high stability of the measure upon resampling.
   - Reliability testing indicates the measure can be used to meaningfully distinguish performance between states.

   Submission document: Testing attachment, section 2a2.3

8. 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.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

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

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)

10. **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** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**
    - Reliability is based on between versus within state variation exclusively.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.
    - The exclusion "beneficiaries who entered hospice" was not tested since number was negligible.
    - The exclusion "discharges followed by transfer or admission to any inpatient or residential setting in 7 or 14 days" could not be tested because the measure cannot be calculated without this exclusion, though these might instead be considered "failures" that add to the denominator. Some quantification regarding this exclusion should be considered.
    - Only 13 states were included in the testing sample, though they seem to be a broad sample based on Medicaid enrollment size.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.
    - No concerns.
    - The measure rates across the 13 states cover a wide range with meaningful variation.
        - The measure rates ranged from 8.9 % to 41.0 % and 24.2 % to 51.1 % continuing care, for 7-day and 14-day follow-ups, respectively.
    - There was significant variation in continuity of care rates between subgroups.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.
    N/A

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.
    - The developer found a relatively small amount of missing data using MAX Validation Reports.
        - The majority of behavioral health encounters from states with a large proportion of their beneficiaries enrolled in a BHO are included in MAX data.

- States will likely have less missing data because they will be able to account for their state-specific codes.
- Is there any concern that services of this type will not be entered into claims?

16. **Risk Adjustment**

16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

16d.**Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

- This measure is a process measure that is not risk adjusted; however, the stratification by various risk factors, including social risk factors, revealed significant differences between subgroups.

VALIDITY: TESTING

17. **Validity testing level:** ☒ **Measure score** ☐ **Data element** ☐ **Both**

18. **Method of establishing validity of the measure score:**

☒ **Face validity**

☒ **Empirical validity testing of the measure score**

☐ **N/A (score-level testing not conducted)**

19. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

- Face Validity tested using 12 member Technical Expert Panel (TEP)
- Convergent Validity was calculated using Spearman rank correlation with two measures of similar concepts:
  - NQF 3312: Percentage of discharges from a detoxification episode for adult Medicaid Beneficiaries, age 18 to 64, that were followed by a treatment service for substance use disorder

- NQF 0576: Percentage of discharges for members 6 years of age and older who were hospitalized for treatment of selected mental illness diagnoses and who had an outpatient visit, an intensive outpatient encounter or partial hospitalization with a mental health practitioner within 7 days after discharge

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

Face Validity

- Twelve out of 12 respondents agreed or strongly agreed that performance scores on the measure "Continuity of care after inpatient or residential treatment for substance use disorder" can be used to distinguish good from poor quality of care.
- Two TEP members noted their strong assessment only applied to the 7-day and not the 14-day continuity rate.

Empirical validity testing of the measure score

- Convergent Validity was not significant, but in the direction posited (see below)

measure	External measure	Spearman rank correlation	95% CI
7-day	NQF 3312 7-day	0.40	(-0.26 ,0.81)
7-day	NQF 0576 7-day	0.34	(-0.29 ,0.76)
14-day	NQF 3312 14 day	0.46	(-0.19 ,0.83)

- Performance rates by state were also provided. States E and L have fairly normal rates for the 14-day measure, but appear to be outliers for NQF 3312 14-day measure.
- Spearman rank correlation coefficients show a non-significant, positive correlation between SUD-18 and #3312 and #0576.
    - Small sample size (n= 13 states)

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

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22. **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.*

Submission document*: Testing attachment, section 2b1.*

☐ **Yes**

☐ **No**

☒ **Not applicable** (data element testing was not performed)

23. **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** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**
    - Face validity was high and convergent validity was marginal.
    - Measure scores show significant variation between states.

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**
    - Is there a concern that two TEP members only supported the 7-day continuity time period?
    - It might be helpful to assess 7-day and 14-day performance scores against one another.
    - Is there concern that the new measure does not significantly correlate with previous related measures?

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

2a1. Reliability – Specifications
Comments:

**Reliability is fine.

**Would argue that follow up should also include PCP. Tere may be hybryd med/psych practices that follow-up is not recorded.

**No concerns would be reliable.

**Data elements are defined, I would like to see a list of prescriptions for SUD.

2a2. Reliability – Testing
Comments:

**One question I have is the availablity of other types of community-based care (AA) and the degree to which regular attendance following intensive SUD care facilitates maintenance of sobriety. This type of community based care is not included in the denominator but could be equally as supportive.

**No. signal to noise ratio was appropriate.

**No.

**No.

**No.

2b1. Validity –Testing

2b4-7. Threats to Validity
2b4. Meaningful Differences
Comments:

**Validity is moderate.

**The measures used to examine convergent validity didn't seem optimal however, prior research and face validity supports the measures's validity.

**No.

**No.

**I am a bit concerned that pharmacotherapy has the same weight as an outpatient visit or telehealth encounter. How is compliance in taking the medication measured? Does a telephone call count as telehealth encounter.

** Validity is questionable in terms of excluding those who return to inpatient care within 7 to 14 days. As reviewers stated, this relapse could be considered a failure of the continuity of care plan.

**None.

** Liited, and Folks may do better simply having follow-up and does not need to BH. I would encouraged any medical follow-up counting.

** No concerns.

** No concerns.

2b2-3. Other Threats to Validity
2b2. Exclusions
2b3. Risk Adjustment
Comments:

** This is a good measure conceptually given the poor continuity of care that those receiving SUD treatment receive.

**None.

**None ID'ed.

**Could we get more information on why rural had a higher rate. Would seem access in rural areas would be harder than in urban areas.

**No concerns.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

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.

- All data elements are in defined fields in electronic claims.

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Is the data collection strategy ready to be put into operational use?

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

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility
Comments:

**Claims data should make this moderately feasible, though it is highly likely diferent states use different codes that could make the data collection strategy challenging.

**No concerns. It's a claims based measure.

**No issues.

**No concerns.

**No concerns.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

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

- Planned Use: Regulatory and Accreditation Programs and Quality Improvement
    - The developer provides a plan for implementation.
        - CMS is developing measures to improve the quality of care of the following Medicaid populations served by CMS's Innovation Accelerator Program.
        - This measure is intended for voluntary use by states to monitor and improve the quality of care provided for Medicaid beneficiaries with substance use disorders.

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 can the performance results be used to further the goal of high-quality, efficient healthcare?
- Is the implementation plan provided appropriate?

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

RATIONALE:

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

- Adoption of this performance measure has the potential to improve the quality of care for Medicaid beneficiaries, who are discharged from inpatient or residential treatment for SUD.
- Low rates indicate room for improvement, and this measure may encourage states to put interventions in place to increase the rates, leading to better outcomes for patients.

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

- There were no unexpected findings identified during testing.

Potential harms

- There were no unexpected harms identified during testing.
- Potential unintended consequences noted in the Evidence section of the submission include: facilities may be incentivized to find any placement quickly, facilities could avoid clients they consider less likely or difficult to achieve continuity of care, states may hold treatment facilities accountable, although continuity also depends on clients' resources and ability to follow through, and agencies that offer multiple levels of care may have an easier time finding post-discharge placement.

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 will not be publicly available. Ideally this would be coupled with a measure for brief counseling for those recently released from SUD treatment. Feedback has been incorporated.
**Similar measures have been used for years for mental health post discharge follow-up so there should be no issue with feasibility.

**Rural under-served area will do worse, and for public reporting could be an issue.
**Had appropriate feedback.
**No concerns.

4b1. Usability – Improvement
Comments:

** No unintended consequences, little harm. but unclear if the benefit is long-term.

**Measure can encourage better transitions of care.

**Again the unintented conseguences have not been throughly studied.

**No issues.

**I don't see any unintended consequences.

Criterion 5: Related and Competing Measures

Related or competing measures
Related measures
0004: Initiation and Engagement of Alcohol and Other Drug Abuse or Dependence Treatment
0576: Follow-Up After Hospitalization for Mental Illness (FUH)
1937: Follow-Up After Hospitalization for Schizophrenia (7- and 30-day)
2605: Follow-Up After Emergency Department Visit for Mental Illness or Alcohol and Other Drug Abuse or Dependence
3312: Continuity of Care for Medicaid Beneficiaries after Detoxification (Detox) From Alcohol and/or Drugs

Harmonization
Numerator:
Timing of continuity of care and population

- The proposed measure specifies continuity of care within 7- and 14-days of discharge and is harmonized with NQF 3312, Continuity of care for Medicaid beneficiaries after detoxification (detox) from alcohol and/or drugs, which also focuses on a SUD population.
- NQF 0576, 1937, and 2605 all specify follow-up within 7 and 30 days.
    - The populations for NQF 0576 and 1937 include patients with mental health related diagnoses rather than focusing on substance use disorders.
    - NQF 2605 has a target mixed population of mental health and SUD patients.
    - In measure testing, stakeholders expressed concern that 30 days is too long for SUD patients to wait for a continuity of care service after discharge from inpatient or residential care.
- NQF 0004 is partially harmonized with the proposed measure in that the initiation visit is specified as within 14 days of the index episode start date (diagnosis).

Diagnoses in the continuity of care visit

- The proposed measure is harmonized with NQF 3312 and NQF 0004 by allowing SUD to either be the primary or a secondary diagnosis for treatment services that count toward continuity in the numerator.
    - Addresses potential inaccuracies in how SUD diagnoses are coded
- NQF 2605 does not allow a secondary SUD diagnosis.
- NQF 0576 and NQF 1937 are not clear on whether only a primary diagnosis is allowed in the numerator.

Services included as continuity of care

- The proposed measure includes pharmacotherapy and telehealth as services that count as continuity of care.
    - Consistent with recent changes made to the 2018 HEDIS specification of NQF 0004.
- NQF 2605, 0576, and 1937 do not include these services.

Practitioners valid for providing follow-up services

- The proposed measure and NQF 2605 allow any practitioner to provide follow-up services
    - Because of expectation that the follow-up services may be provided by primary care clinicians.
- NQF 0576 and 1937 only allow non-mental health practitioners in specified settings and with specific diagnosis codes.

Denominator:
Diagnoses in denominator

- The denominators for the proposed measure and all the related measures are harmonized in requiring a primary diagnosis for the condition that is the measure's focus.

Age

- The proposed measure is intended for an adult Medicaid population and excludes adults over 64 years, because complete data on services received by dually-eligible (Medicaid and Medicare) adults are not available in Medicaid data.
- Similar to NQF 3312 and NQF 1937, it includes ages 18-64.
- NQF 2605 includes adults age 18 and older.
- NQF 0576 includes individuals age 6 and older.
- NQF 0004 includes age 13 and older.

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

5. Related and Competing

Comments

**Other similar measures extend to 30 days and I think this one should as well.

**No.

**Hedis- has measure for inpatient.

**This measure seems to be similar to #3312 except that #3453 includes telehealth (which is great) so the question I have is do we need two different measures that are only different due to telehealth being included? Can we keep this and retire #3312. I think it is difficult in the filed for people to know that is the only difference. Thanks.

**Unaware of any.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 01/22/2019

There have been no comments or support/non-support choices as of this date.

Developer Submission

Brief Measure Information

NQF #: 3453

Corresponding Measures:

De.2. Measure Title: Continuity of care after inpatient or residential treatment for substance use disorder (SUD)

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

De.3. Brief Description of Measure: Percentage of discharges from an inpatient or residential treatment for substance use disorder (SUD) for Medicaid beneficiaries, ages 18 to 64, which was followed by a treatment service for SUD. SUD treatment includes having an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth encounter, or filling a prescription or being administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

1b.1. Developer Rationale: Continuity of care helps to sustain the gains attained in initial treatment and to prevent relapses. There is general agreement that continuity of care (including encounters with the health system within a defined period of time) after discharge from inpatient or residential care for SUD is related to better outcomes including reduced substance use (DeMarce, Lash, Stephens, Grambow, & Burden, 2008; McKay & Hiller-Sturmhofel, 2011), readmissions (Mark et al., 2013; Reif et al., 2017), and criminal justice involvement (McKay, 2009), lower risk of death in the two post-discharge years (Harris et al., 2015), and improved employment status (McKay, 2009).

Although the definition of continuity varies across studies (including outpatient treatment, inpatient or residential treatment, pharmacotherapy), the findings are consistent. In spite of the benefits of continuity of care, many patients do not receive treatment services after being discharged from residential or inpatient hospital care (Garnick, Lee, Horgan, Acevedo, & Washington Circle Public Sector, 2009; Harris et al., 2006; Rubinsky et al., 2017). Continuity after inpatient or residential treatment has been found to be generally low and the variation in continuity rates suggests that there is substantial opportunity for improvement.

- A study using claims data for Medicaid enrollees ages 18-64 who had an inpatient hospital or residential detoxification admission for SUD found that 75% of these enrollees did not receive at least one continuity of care services (outpatient, intensive outpatient, or residential) within 14 days of discharge. This varies by state and ranged from 50% to 83% (Reif, Acevedo, Garnick, & Fullerton, 2017).
- In five states' public sector SUD treatment systems, continuity of care rates for receiving at least one residential or outpatient treatment service within 14 days after discharge, ranged from 15% to 60% (Garnick et al., 2009).
- In a study with veterans, continuity of care was defined as engaging in continuing care, i.e., the total number of consecutive months after intensive treatment (e.g., inpatient or residential) during which the patient had 2 or more clinic visits and no inpatient SUD or psychiatric readmissions. Only 32% of patients had two or more continuity of care visits during the month

after discharge. At three months after discharge from inpatient/residential only 17% of patients had continuity care and this was further reduced to 10% at six months after discharge (Schaefer, Ingudomnukul, Harris, & Cronkite, 2005).

- Another study found continuity of care after residential treatment to range from 4% to 91% for SUD care within 30 days of discharge, among Veterans Health Administration (VHA) programs (Rubinsky et al., 2017). Continuity of care was defined in this study as having at least one outpatient SUD or mental health visit within 30 days. SUD and mental health "visits" were defined as clinical encounters within a SUD or mental health clinic, respectively, and included individual and group encounters as well as teleheath care.
- Yet another study of VHA residential treatment programs found that more than half (59%) of patients did not engage in two or more outpatient SUD treatment visits during the first month after discharge (Harris et al., 2006).

Use of a performance measure to identify patients who are less likely to have continuity of care or have shorter stays in follow-up care is critical in targeting extra efforts in engaging these patients (Harris, McKellar, Moos, Schaefer, & Cronkite, 2006).

S.4. Numerator Statement: Discharges in the denominator with an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth or filled a prescription for or were administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

S.6. Denominator Statement: Adult Medicaid beneficiary discharges from inpatient or residential treatment for SUD with a principal diagnosis of SUD during from January 1 to December 15 of the measurement year.

S.8. Denominator Exclusions: Exclude from the denominator for both rates:

- Discharges with hospice services during the measurement year
- Both the initial discharge and the admission/direct transfer discharge if the admission/direct transfer discharge occurs after December 15 of the measurement year.

Discharges followed by admission or direct transfer to inpatient or SUD residential treatment setting within 7 or 14-day continuity of care period. These discharges are excluded from the measure because transfer, hospitalization or admission to residential treatment within 7 or 14 days may prevent a continuity of care visit from taking place. An exception is admission to residential treatment following discharge from inpatient treatment; we do not exclude these admissions, because continuity into residential treatment after inpatient treatment is considered appropriate treatment.

De.1. Measure Type: Process

S.17. Data Source: Claims

S.20. Level of Analysis: Population : Regional and State

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:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not applicable; this measure is not a paired or grouped measure.

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

4.SUD-18_Evidence_Attachment_FINAL_SUD_team_09_04.18.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*)**:** NA

Measure Title: Continuity of care after inpatient or residential treatment for substance use disorder (SUD)

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

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:

☐ 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.

Despite literature showing the beneficial effects of continuity of care after inpatient or residential treatment for substance use disorder (SUD), continuity rates are low, leaving much room for improvement (Garnick, Lee, Horgan, Acevedo, & Washington Circle Public Sector, 2009; Harris, McKellar, Moos, Schaefer, & Cronkite, 2006; Rubinsky et al., 2017). Endorsement and implementation of a measure for quality improvement (internal monitoring and external with benchmarking) may improve the rate of follow-up care post-inpatient or residential treatment for SUD. The logic model shows potential benefits of such a measure at the client level that includes reduction in hospital admissions (Blodgett, Maisel, Fuh, Wilbourne, & Finney, 2014; Mark et al., 2013; McKay, 2009; Reif, Acevedo, Garnick, & Fullerton, 2017), substance use (DeMarce, Lash, Stephens, Grambow, & Burden, 2008; McKay, 2005), and criminal justice activity (McKay, 2009), as well as improved employment status (McKay, 2009). Continuity of care also helps a client to sustain gains from the initial treatment (Rubinsky et al., 2017). The benefits to society include a reduction in cost related to lower crime rates and lower health care costs (Popovici, French, & McKay, 2008).

The logic model below shows how implementation of the measure can lead to benefits for clients and society in addition to healthcare savings. However, implementation of the measure does have intermediate financial costs that include the costs of coordinating new referrals and costs to Medicaid if clients receive more services that are expected which are expected to be offset by longer term savings. As with any intervention there are also potential unintended consequences, such as facilities may be incentivized to find any placement quickly, at the expense of finding the right placement more slowly. Factors that influence the measure include 1) systems organization and capacity limitations and 2) data availability and completeness. The first relates to capacity of the treatment system and whether there is adequate available treatment for clients discharged from inpatient or residential care. The second relates to the quality of data that reporting entities have available to them for calculating the measure and variations in state Medicaid coverage of SUD treatment services.

Measure Information

Measure description: Percentage of discharges from an inpatient or residential treatment for substance use disorder (SUD) for Medicaid beneficiaries, ages 18 to 64, which was followed by a treatment service for SUD. SUD treatment includes having an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth encounter, or filling a prescription or being administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

Numerator: Discharges in the denominator with an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth encounter or filled a prescription for or were administered or ordered a medication for SUD.

Denominator: Adult Medicaid beneficiary discharges from inpatient or residential treatment for SUD with a principal diagnosis of SUD during from January 1 to December 15 of the measurement year.

Exclusions:

- Discharges with hospice services during the measurement year
- The initial discharge and the admission/direct transfer discharge if the admission/direct transfer discharge occurs after December 15 of the measurement year.
- Discharges followed by admission or direct transfer to inpatient or SUD residential treatment settings within 7 or 14-day continuity of care period. These discharges are excluded from the measure because transfer, hospitalization or admission to residential treatment within 7 or 14 days may prevent a continuity of care visit from taking place. An exception is admission to residential treatment following discharge from inpatient treatment.

Measure Implementation

Quality improvement (internal to a specific organization)—supports internal quality monitoring and improvement at the state, program, or provider level

Quality improvement with benchmarking (external benchmarking to multiple organizations)

Benefits

Health care:

- Continuity of care after inpatient or residential treatment is generally low and varies; thus, there is much room for improvement in continuity of care

Health outcomes after inpatient or residential discharge:

Impact on clients

- Fewer hospital readmissions
- Less substance use and relapse
- Improved employment status
- Reduction in criminal justice activity

Impact on society

- Reduction in costs related to lower crime rates and lower health care costs

Health care/Medicaid savings:

- Lower costs as a result of continuity of care that helps to sustain a patient's gains from the initial treatment and to prevent relapses

Costs and Unintended Consequences

Implementation costs:

- Low cost to adopt measure because it relies on administrative data
- Time for staff to add measure to their current set of measures
- Cost for programmers to review specifications and add coding to current programs
- Cost for training and hiring additional staff to oversee transition of clients to next level of care
- Cost for program to coordinate referrals and transport clients to next level of care and develop network of referral partners
- Expansion of system capacity so there are enough treatment slots to allow timely continuity of care.

Intervention costs:

- Increased cost to Medicaid if more people receive at least one continuity of care service
- Increased cost to client for additional treatment, lost days of work to enter treatment, transportation costs, and possible child care costs

Unintended consequences:

- If there is insufficient capacity, facilities may avoid clients whom they consider less likely to have continuity.
- To meet the measure's requirement of continuity of care within 7 or 14 days, providers may pay less attention to finding an appropriate placement
- Programs may be held accountable for events that are beyond their control (e.g., client motivation or system capacity or organization)

Influencing Factors

- Health system issues
  - System organization and capacity
  - Location of necessary services
- Data availability and completeness
  - States' coverage of SUD treatment services
  - Organization of Medicaid systems

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

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

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.

The evidence on which this measure is based is the result of a targeted review of the literature on the benefits of continuity of care (including any encounter with the health system within a defined period of time) after discharge from inpatient or residential care for SUD and the extent to which this occurs.

Note that in the literature, the terms "continuity of care" and "follow-up" often are used interchangeably since both refer to the same concept of receiving continuing care after an initial service for SUD.

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

Continuity of care (including any encounter with the health system within a defined period of time) is widely recommended after inpatient or residential SUD treatment, although the studies vary in terms of definitions of continuity (timeframe and requirements for numbers and types of treatment) and outcomes (mortality, substance use, readmission, retention in treatment, problem severity, employment, and criminal justice involvement) (Blodgett et al., 2014; McKay, 2005, 2009; McKay & Hiller-Sturmhofel, 2011). Review articles have reported that receiving more types of services, longer duration, or more frequent visits during continuity of care treatment is related to better outcomes (Blodgett et al., 2014; McKay, 2009; McKay & Hiller-Sturmhofel, 2011).

In addition to review articles, specific articles also report a range of outcomes related to continuity of care after inpatient or residential treatment for SUD, including mortality, readmission, retention in treatment and substance use; and that timely continuity of care shortly after discharge from initial treatment is good care (Table 1):

Mortality. One study addressed mortality, showing that continuity of care is related to lower risk of death in the two post-discharge years. Patients who had continuity of care within 14 days of discharge from residential SUD treatment were also found to have a lower two-year mortality ratio, (OR = 0.77, p<.008) (Harris et al., 2015).

Readmission. One study (Reif et al., 2017) with a sample of inpatient hospital admissions among Medicaid beneficiaries with a substance use disorder or residential detoxification admission found that receipt of residential treatment after discharge, compared to no follow-up, was related to a lower hazard of readmission (HR=0.50, p<.001). Furthermore, medication-assisted treatment for SUD after an inpatient or detox stay helps reduce the likelihood of readmissions (HR= 0.61, p<.001). In another study of Medicaid beneficiaries who had inpatient treatment for mental health or SUDs, continuity of care treatment at community mental health centers was associated with a lower probability of readmission (Mark et al., 2013). This study examined a

facility-level variable, the percent of patients at the hospital who received post-discharge follow-up at a community mental health center within 7 days, and the influence on patients being readmitted. Findings showed that for every 1 percent increase in patients receiving follow-up care, there was a five percent reduction in the probability of readmission.

Retention in Treatment. Patients that were randomized to receive continuing care adherence intervention had better retention through 3-month follow-up and a higher abstinence rate at 1-year follow-up, as compared to patients receiving standard treatment that did not provide interventions to support adherence to continuity of care. The continuing care adherence intervention consisted of contracting, prompting, and reinforcing attendance (DeMarce et al., 2008).

Reduced Substance Use or Abstinence. Adolescents treated in publicly funded long-term residential care were more likely to be abstinent at follow-up when they met with a clinician to initialized continuity of care services as soon as possible after discharge compared to those who received standard care (treatment facility did not follow-up to see whether the patient initiated continuity of care services) (Garner, Godley, Funk, Lee, & Garnick, 2010).

Table 1. Studies of Continuity Impact on Mortality, Substance Use, Retention in Treatment and Readmission

Study	Design	Time After Discharge to Continuity of Care	Outcomes	Results
Harris et al. 2015	Veterans Health Administration administrative data; cross sectional analysis (FY 2009), using propensity score weighted mixed effects regression models	Within 14-days after residential discharge	2-year mortality	76% of 10,064 patients had a continuity of care service within 14 days of discharge. Patients who had continuity of care had a lower 2-year mortality rate (OR= 0.77, p=0.008) but no difference in subsequent detoxification episodes within six months relative to patients without a follow-up visit.
Reif et al. 2017	Medicaid Analytic eXtract claims data (2008; N=30,439) for ten states; cross sectional analysis, survival analysis (Cox proportional hazards) to estimate the effect of follow-up services and MAT on time to a behavioral health readmission with censoring at 90 days	Within 14-days after discharge from inpatient hospital care or residential detoxification	Readmissions after inpatient hospital admission for SUD or after residential detoxification	Medication assisted treatment and residential treatment provided after discharge were associated with reduced risk of 90-day behavioral health readmission (HR= 0.05, p<.001 and HR= 0.61, p<.001 for residential and MAT, respectively.

Study	Design	Time After Discharge to Continuity of Care	Outcomes	Results
Garner et al. 2010	Adolescents in publicly funded long-term residential treatment for at least 7 days (N=342) were randomized to either a control group that received standard care (SC) or to one of three assertive care conditions (ACC): assertive continuing care, Contingency management, or both.	Within 14-days of discharge from long-term residential care	Substance use and substance use problems as measure by the GAIN substance use scale (SPS). Recovery status measured as abstinence and no dependence symptoms while living in the community during past 30 days prior to follow-up.	Continuity of care was higher for those in the ACC (78%) compared with SC (56%) and continuity of care was a significant predictor of 3-month recovery status (OR= 1.92, p<.05).
Mark et al. 2013	Analyses conducted with MarketScan® Multistate Medicaid Database, 2004-2009, cross section multivariate analyses controlling for hospital and patient level characteristics.	Within 7-days after discharge	Readmissions to treatment with primary SUD diagnosis in days 8 to 30 days after discharge	The hospitals with a larger percentage of patients with a community mental health center (CMHC) follow up were associated with lower probability of readmission. A 1 percent increase in a hospital's percent of patients receiving post-discharge continuity of care in a CMHS within 7 days was associated with a 5 percent reduction in the probability of being readmitted.

Study	Design	Time After Discharge to Continuity of Care	Outcomes	Results
DeMarce et al. 2008	Randomized controlled study, 150 participants randomized to receive behavioral continuing care adherence intervention or standard treatment. The intervention included meeting with therapist during last week of residential care to develop a behavioral continuing care contract, and meeting again 9 weeks after discharge. The intervention group received attendance prompts, feedback prior to aftercare therapy sessions, certificates for treatment attendance, and AA/NA meetings.	Beginning aftercare after residential treatment and completing 3-, 7-, and 12- month treatment. Beginning aftercare defined as attending at least two treatment sessions per month.	Retention/length of stay in aftercare, abstinence	Continuity of care supports related to longer stay in aftercare and higher 1-year abstinence rate

Measure impact on healthcare costs

Continuity of care helps to sustain a patient's gains from the initial treatment and to prevent relapse. In a review of economic evaluations of continuing care studies, Popovici et al. (2008) found that the total benefits generated by continuity of care interventions, including those services after inpatient admissions, exceeded the total medical costs.

Given the impact of continuity of care following inpatient or specialty residential treatment for SUDs on readmission rates, greater performance on the measure has the potential to reduce health care costs. One study of Medicaid beneficiaries in 2004–2009 offers insights into the gross impact of such improvements on costs: Mark et al. (2013) examined the median rate of behavioral health-related hospital readmissions within 8 to 30 days after discharge from another inpatient stay with a principal mental health or SUD diagnosis. The findings from this study, along with an estimate of the cost of inpatient treatment, enable us to approximate the impact of continuity of care on health care costs. Mark et al. (2013) reported that 11 percent of the 121,271 discharges (13,340) had a behavioral health readmission. Given a cost of $6,300 for inpatient treatment for mental health or SUD (Heslin, Elixhauser, & Steiner, 2015), multiplying 13,340 readmission times $6,300 per inpatient treatment is $84,042,000. For each percentage-point decrease in behavioral health readmissions (a decrease of 1,213 discharges multiplied by $6,300 per inpatient treatment), there is a savings of $7,641,900. Note that Reif et al. (2017) reported even higher readmission rates with a principal behavioral health diagnosis (29.3 percent within 90 days). Thus, the total costs derived from that study—and the potential savings—would be even higher.

These estimates should be viewed with caution because of the age of the data, the fact that the inpatient cost estimate is based on all payers rather than on Medicaid only, the inclusion of both mental health and SUD

treatment, and uncertainty about the magnitude of change in readmission rates associated with increased continuity of care. In addition, only including the costs of averted inpatient treatment may lead us to underestimate the total savings. This is because continuity of care may result in better long-term outcomes and avert costs of additional episodes of SUD treatment. Despite these caveats, we could potentially see large savings from an improvement in continuity of care, which is associated with reductions in readmissions, because of the high cost of inpatient treatment.

Performance Gap

In spite of the benefits of continuity of care, many patients do not receive additional treatment services after being discharged from residential or inpatient care for SUD (Garnick et al., 2009; Harris et al., 2006; Rubinsky et al., 2017). In summary, the studies described below, examining continuity of care within 14 days of inpatient discharge, found the rate to be about 25 to 47 percent, and rates of continuity of care after residential treatment to range from 15 to 60 percent. Continuity of care after inpatient or residential treatment is generally low, and the variation in continuity rates suggests that there is room for improvement.

- Using claims data for Medicaid enrollees ages 18 to 64 who had an inpatient hospital or residential detox admission for SUD, researchers found that 75 percent of these enrollees received no continuity of care services (outpatient or pharmacotherapy) within 14 days of discharge. This varied by state and ranged from 50 to 83 percent (Reif et al., 2017).
- In a study of five states' public-sector SUD treatment systems, continuity of care rates within 14 days of residential discharge were 15 to 60 percent, and continuity rates after an inpatient hospital stay were 27 to 47 percent (Garnick et al., 2009).
- One study showed that among 63 residential treatment programs funded by the Veterans Health Administration (VHA) to treat SUD, there was a mean rate of 63 percent continuity of care within 30 days of discharge (Rubinsky et al., 2017). Continuity of care was defined in this study as at least one outpatient SUD or mental health visit within 30 days. SUD and mental health "visits" were defined as clinical encounters within a SUD or mental health clinic, respectively, and included individual and group encounters as well as telehealth care.
- Another study of VHA residential treatment programs showed that more than half (59 percent) of patients did not have two or more outpatient SUD treatment visits during the first month after discharge (Harris et al., 2006).

Potential Unintended Consequences of the Measure

The implementation of the SUD-18 measure may have the following potential untoward effects:

- If there is insufficient capacity in the treatment system, facilities could avoid clients they consider less likely or difficult to achieve continuity of care, such as people who are homeless.
- To meet the measure's requirements of a treatment service within 7 or 14 days, providers may place beneficiaries in less than optimally matched locations if those are the only ones available.
- Under this measure of care coordination, states may hold treatment facilities accountable, although continuity also depends on clients' resources and ability to follow through.
- Inpatient or residential treatment facilities may work with agencies that offer multiple levels of care. Such agencies, even if they are not co-located but are managed under the same organizational umbrella, may have an easier time finding post-discharge placements compared to agencies that offer only one level of care.

Net benefit

Overall, the evidence suggests that the potential benefits of implementing SUD-18 outweigh the potential costs or unintended consequences. Benefits to Medicaid beneficiaries who have a SUD and receive follow-up within 7 or 14 days following discharge from inpatient or residential settings include reduced substance use and relapse, fewer readmissions to inpatient hospitals, less involvement in criminal justice, and improved

employment status. Moreover, the benefits to society include lower costs related to criminal activity and health care. Health care costs for Medicaid could be reduced to the extent that continuity of care helps sustain beneficiaries' gains from the initial treatment and prevents readmissions.

Because the measure is based on claims data, the cost to adopt it is relatively low. But generating improvement in the areas covered by the measure will require additional efforts by facilities to help clients pursue continuity of care and to review the treatment system's capacity to provide this care. Moreover, any improvement in the SUD-18 measure implies a cost to Medicaid for the continuity of care services. However, the benefits of continuity of care to Medicaid beneficiaries and to society, including those described previously, outweigh the costs of implementing and using this measure.

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

PubMed searches were conducted using keywords: continuity of care, follow-up treatment, residential, inpatient, substance use disorder, treatment, for any type of study since 2000. We focused on the extent to which continuity of care occurs after discharge and the benefits of receiving additional care after leaving inpatient and residential care for SUD.

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

Blodgett, J. C., Maisel, N. C., Fuh, I. L., Wilbourne, P. L., & Finney, J. W. (2014). How effective is continuing care for substance use disorders? A meta-analytic review. *J Subst Abuse Treat, 46*(2), 87-97. doi: 10.1016/j.jsat.2013.08.022

DeMarce, J. M., Lash, S. J., Stephens, R. S., Grambow, S. C., & Burden, J. L. (2008). Promoting continuing care adherence among substance abusers with co-occurring psychiatric disorders following residential treatment. *Addict Behav, 33*(9), 1104-1112. doi: 10.1016/j.addbeh.2008.02.008

Garner, B. R., Godley, M. D., Funk, R. R., Lee, M. T., & Garnick, D. W. (2010). The Washington Circle continuity of care performance measure: predictive validity with adolescents discharged from residential treatment. *J Subst Abuse Treat, 38*(1), 3-11. doi: 10.1016/j.jsat.2009.05.008

Garnick, D. W., Lee, M. T., Horgan, C. M., Acevedo, A., & Washington Circle Public Sector, Workgroup. (2009). Adapting Washington Circle performance measures for public sector substance abuse treatment systems. *J Subst Abuse Treat, 36*(3), 265-277. doi: 10.1016/j.jsat.2008.06.008

Harris, A. H., Gupta, S., Bowe, T., Ellerbe, L. S., Phelps, T. E., Rubinsky, A. D., . . . Trafton, J. (2015). Predictive validity of two process-of-care quality measures for residential substance use disorder treatment. *Addict Sci Clin Pract, 10*, 22. doi: 10.1186/s13722-015-0042-5

Harris, A. H., McKellar, J. D., Moos, R. H., Schaefer, J. A., & Cronkite, R. C. (2006). Predictors of engagement in continuing care following residential substance use disorder treatment. *Drug Alcohol Depend, 84*(1), 93-101. doi: 10.1016/j.drugalcdep.2005.12.010

Heslin, K.C., Elixhauser, A., & Steiner, C.A. (2015). Hospitalizations Involving Mental and Substance Use Disorders Among Adults, 2012. *HCUP Statistical Brief #191. June 2015. Agency for Healthcare Research and Quality, Rockville, MD.* Retrieved August 6, 2018, from http://www.hcup-us.ahrq.gov/reports/statbriefs/sb191-Hospitalization-Mental-Substance-Use-Disorders-2012.pdf

Mark, T. L., Tomic, K. S., Kowlessar, N., Chu, B. C., Vandivort-Warren, R., & Smith, S. (2013). Hospital readmission among medicaid patients with an index hospitalization for mental and/or substance use disorder. *J Behav Health Serv Res, 40*(2), 207-221. doi: 10.1007/s11414-013-9323-5

McKay, J. R. (2005). Is there a case for extended interventions for alcohol and drug use disorders? *Addiction, 100*(11), 1594-1610. doi: 10.1111/j.1360-0443.2005.01208.x

McKay, J. R. (2009). Continuing care research: what we have learned and where we are going. *J Subst Abuse Treat, 36*(2), 131-145. doi: 10.1016/j.jsat.2008.10.004

McKay, J. R., & Hiller-Sturmhofel, S. (2011). Treating alcoholism as a chronic disease: approaches to long-term continuing care. *Alcohol Res Health, 33*(4), 356-370.

Popovici, I., French, M. T., & McKay, J. R. (2008). Economic evaluation of continuing care interventions in the treatment of substance abuse: recommendations for future research. *Eval Rev, 32*(6), 547-568. doi: 10.1177/0193841X08316311

Reif, S., Acevedo, A., Garnick, D. W., & Fullerton, C. A. (2017). Reducing Behavioral Health Inpatient Readmissions for People With Substance Use Disorders: Do Follow-Up Services Matter? *Psychiatr Serv, 68*(8), 810-818. doi: 10.1176/appi.ps.201600339

Rubinsky, A. D., Ellerbe, L. S., Gupta, S., Phelps, T. E., Bowe, T., Burden, J. L., & Harris, A. H. S. (2017). Outpatient Continuing Care after Residential Substance Use Disorder Treatment in the U.S. Veterans Health Administration: Facilitators and Challenges. *Subst Abus*, 0. doi: 10.1080/08897077.2017.1391923

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.

Continuity of care helps to sustain the gains attained in initial treatment and to prevent relapses. There is general agreement that continuity of care (including encounters with the health system within a defined period of time) after discharge from inpatient or residential care for SUD is related to better outcomes including reduced substance use (DeMarce, Lash, Stephens, Grambow, & Burden, 2008; McKay & Hiller-Sturmhofel, 2011), readmissions (Mark et al., 2013; Reif et al., 2017), and criminal justice involvement (McKay, 2009), lower risk of death in the two post-discharge years (Harris et al., 2015), and improved employment status (McKay, 2009).

Although the definition of continuity varies across studies (including outpatient treatment, inpatient or residential treatment, pharmacotherapy), the findings are consistent. In spite of the benefits of continuity of care, many patients do not receive treatment services after being discharged from residential or inpatient hospital care (Garnick, Lee, Horgan, Acevedo, & Washington Circle Public Sector, 2009; Harris et al., 2006; Rubinsky et al., 2017). Continuity after inpatient or residential treatment has been found to be generally low and the variation in continuity rates suggests that there is substantial opportunity for improvement.

- A study using claims data for Medicaid enrollees ages 18-64 who had an inpatient hospital or residential detoxification admission for SUD found that 75% of these enrollees did not receive at least one continuity of care services (outpatient, intensive outpatient, or residential) within 14 days of discharge. This varies by state and ranged from 50% to 83% (Reif, Acevedo, Garnick, & Fullerton, 2017).
- In five states' public sector SUD treatment systems, continuity of care rates for receiving at least one residential or outpatient treatment service within 14 days after discharge, ranged from 15% to 60% (Garnick et al., 2009).
- In a study with veterans, continuity of care was defined as engaging in continuing care, i.e., the total number of consecutive months after intensive treatment (e.g., inpatient or residential) during which the patient had 2 or more clinic visits and no inpatient SUD or psychiatric readmissions. Only 32% of patients had two or more continuity of care visits during the month after discharge. At three months after discharge from inpatient/residential only 17% of patients had continuity care and this was further reduced to 10% at six months after discharge (Schaefer, Ingudomnukul, Harris, & Cronkite, 2005).

- Another study found continuity of care after residential treatment to range from 4% to 91% for SUD care within 30 days of discharge, among Veterans Health Administration (VHA) programs (Rubinsky et al., 2017). Continuity of care was defined in this study as having at least one outpatient SUD or mental health visit within 30 days. SUD and mental health "visits" were defined as clinical encounters within a SUD or mental health clinic, respectively, and included individual and group encounters as well as teleheath care.
- Yet another study of VHA residential treatment programs found that more than half (59%) of patients did not engage in two or more outpatient SUD treatment visits during the first month after discharge (Harris et al., 2006).

Use of a performance measure to identify patients who are less likely to have continuity of care or have shorter stays in follow-up care is critical in targeting extra efforts in engaging these patients (Harris, McKellar, Moos, Schaefer, & Cronkite, 2006).

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.*

We used 2014 Medicaid data to test the measure in 13 states for two rates, 7-day and 14-day continuity of care. State names are redacted. The overall 7-day continuity of care rate, across all states tested, was 18.4% and ranged from 8.9% in State D to 41.0% in State L. The overall 14-day continuity of care rate was slightly higher at 24.2% since the window of time to have a continuity of care service was extended. The range for the 14-day rate was 13.2% in State D to 51.1% in State L.

Below we present the performance rates, overall and for each state, for the 7-day and 14-day continuity of care measure.

7 day continuity of care results:
Overall rate across all states:
Numerator: 15,508
Denominator: 84,375
Rate: 18.4%
State A:
Numerator: 863
Denominator: 4186
Rate: 20.6%
95% CI = (19.39, 21.84)
State B:
Numerator: 130
Denominator: 1178
Rate: 11.0%
95% CI = (9.25, 12.82)
State C:
Numerator: 869
Denominator: 2919
Rate: 29.8%
95% CI = (28.11, 31.43)
State D:

Numerator: 181
Denominator: 2036
Rate: 8.9%
95% CI = (7.65, 10.13)
State E:
Numerator: 1114
Denominator: 8220
Rate: 13.6%
95% CI = (12.81, 14.29)
State F:
Numerator: 235
Denominator: 1077
Rate: 21.8%
95% CI = (19.35, 24.29)
State G:
Numerator: 188
Denominator: 1073
Rate: 17.5%
95% CI = (15.25, 19.80)
State H:
Numerator: 5932
Denominator: 39,123
Rate: 15.2%
95% CI = (14.81, 15.52)
State I:
Numerator: 1746
Denominator: 7516
Rate: 23.2%
95% CI = (22.28, 24.19)
State J:
Numerator: 2600
Denominator: 12,183
Rate: 21.3%
95% CI = (20.61, 22.07)
State K:
Numerator: 407
Denominator: 1594
Rate: 25.5%
95% CI = (23.39, 27.67)
State L:
Numerator: 934
Denominator: 2276
Rate: 41.0%
95% CI = (39.02, 43.06)

State M:
Numerator: 309
Denominator: 994
Rate: 31.1%
95% CI = (28.21, 33.96)
14-day continuity of care results:
Overall rate across all states:
Numerator: 18,745
Denominator: 77,635
Rate: 24.1%
State A:
Numerator: 1014
Denominator: 3423
Rate: 29.6%
95% CI = (28.09, 31.15)
State B:
Numerator: 169
Denominator: 1117
Rate: 15.1%
95% CI = (13.03, 17.23)
State C:
Numerator: 1074
Denominator: 2783
Rate: 38.6%
95% CI = (36.78, 40.40)
State D:
Numerator: 255
Denominator: 1937
Rate: 13.2%
95% CI = (11.66, 14.67)
State E:
Numerator: 1519
Denominator: 7630
Rate: 19.9%
95% CI = (19.01, 20.80)
State F:
Numerator: 265
Denominator: 1016
Rate: 26.1%
95% CI = (23.38, 28.78)
State G:
Numerator: 234
Denominator: 988
Rate: 23.7%

95% CI = (21.03, 26.34)
State H:
Numerator: 6712
Denominator: 35,610
Rate: 18.8%
95% CI = (18.44, 19.25)
State I:
Numerator: 2284
Denominator: 7031
Rate: 32.5%
95% CI = (31.39, 33.58)
State J:
Numerator: 3283
Denominator: 11,451
Rate: 28.7%
95% CI = (27.84, 29.50)
State K:
Numerator: 485
Denominator: 1532
Rate: 31.7%
95% CI = (29.33, 33.99)
State L:
Numerator: 1121
Denominator: 2192
Rate: 51.1%
95% CI = (49.05, 53.23)
State M:
Numerator: 330
Denominator: 925
Rate: 35.7%
95% CI = (32.59, 38.76)

Source: Based on analysis of 2014 Medicaid Alpha-MAX data - eligible (EL), inpatient (IP), other services (OT), long-term care (LT) and drug (RX) files. The other services file contains facility and individual provider services data. Most notably, it may contain both residential and other stayover service claims data as claims are assigned to MAX claims file types based upon the category of service provided.

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.

Not applicable; see performance data above in 1b.2.

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.

We tested a 7- and 14-day measure and results are presented below for the overall population in states included in testing and by Medicaid beneficiary category, age, gender, race/ethnicity, and rural/urban.

Across all states, there were higher continuity rates for Medicaid-aged beneficiaries, females, White beneficiaries and beneficiaries of Other race/ethnicity groups, and those from rural areas, although there were some differences across states.

7-day continuity of care results:

Total:

N=84,375

Rate = 18.4%

Medicaid beneficiary category

Aged

N= 1615

Rate = 38.8%

Blind-disabled

N= 23,851

Rate = 16.2%

Adult, Non-disabled

N= 57,496

Rate = 18.8%

Child

N= 1413

Rate = 14.8%

Age

18-24

N= 8692

Rate = 19.8%

25-44

N= 42,322

Rate = 20.5%

45-64

N= 33,361

Rate = 15.3%

Gender

Female

N= 30,014

Rate = 21.9%

Male

N= 54,361

Rate = 16.5%

Race/Ethnicity

American Indian or Alaskan Native

N= 731

Rate = 19.2%
Asian
N= 1202
Rate = 9.1%
Black
N= 20,754
Rate = 11.3%
Hispanic/Latino
N= 10,536
Rate = 20.6%
Native Hawaiian or Pacific Islander
N= 114
Rate = 17.5%
Other Race
N= 199
Rate = 22.1%
Unknown Race
N= 4116
Rate = 14.9%
White
N= 46,723
Rate = 21.6%
Urbanicity
Rural
N= 9572
Rate = 24.5%
Urban
N= 74,681
Rate = 17.6%
Urban Rural Unknown
N= 122
Rate = 32.8%
14-day continuity of care results:
Total:
N=77,635
Rate = 24.1%
Medicaid beneficiary category
Aged
N= 1545
Rate = 49.0%
Blind-disabled
N= 21,788
Rate = 21.8%
Adult, Non-disabled

N= 52,974
Rate = 24.5%
Child
N= 1328
Rate = 19.1%
Age
18-24
N= 8198
Rate = 25.8%
25-44
N= 39,163
Rate = 26.7%
45-64
N= 30,274
Rate = 20.3%
Gender
Female
N= 28,110
Rate = 28.7%
Male
N= 49,525
Rate = 21.6%
Race/Ethnicity
American Indian or Alaskan Native
N= 681
Rate = 27.8%
Asian
N= 1066
Rate = 11.9%
Black
N= 18,762
Rate = 15.3%
Hispanic/Latino
N= 9495
Rate = 24.9%
Native Hawaiian or Pacific Islander
N= 106
Rate = 24.5%
Other Race
N= 186
Rate = 28.0%
Unknown Race
N= 3898
Rate = 20.1%

White

N= 43,441

Rate = 28.4%

Urbanicity

Rural

N= 9174

Rate = 32.4%

Urban

N= 68,345

Rate = 23.0%

Urban Rural Unknown

N= 116

Rate = 40.5%

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

Not applicable; see data on disparities above in 1b.4.

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.)*

The measure does not yet have published specifications. Therefore no link exists

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)*

Attachment: Attachment SUD-18_measure_value_sets_FINAL_08.09.18_tested_sets_-_locked.xlsx

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.

No

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.

Not applicable. This is a new measure.

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).

Discharges in the denominator with an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth or filled a prescription for or were administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

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).

The measure will report two rates, continuity of care within 7 days and within 14 days after discharge.

The numerator includes discharges with any of the following after inpatient or residential treatment:

- Outpatient visit, intensive outpatient encounter or partial hospitalization with a primary or secondary SUD diagnosis on the day after discharge through day 7 or 14.
- Telehealth encounter for SUD on the day after discharge through day 7 or 14
- Pharmacotherapy (filling a prescription or being administered or ordered a medication) on day of discharge through day 7 or 14
- For inpatient discharges only, residential admissions on day 3 through day 7 or day 14

Public comments supported a measure for 7- and 14-day continuity and voiced that beyond that would be too long, risking losing the patient from the treatment system. The Technical Expert Panel unanimously agreed on the appropriateness of 7-day continuity of care. However, three TEP members felt that 14-days continuity of care is too long. Our approach balances clinical best practice thinking that the sooner the patient is connected to treatment the better while also allowing treatment programs more time for placement of patients in continuing treatment. Because it may be difficult at times for treatment programs to place clients in continuing care in a timely fashion after discharge due to limits in systems capacity, it is particularly important to allow more time for continuity of care to occur.

Inpatient or residential treatment was considered to be SUD related if it had a primary SUD diagnosis or a procedure indicating SUD. SUD diagnoses are identified through ICD-9 codes. Procedures are defined using a

combination of Healthcare Common Procedure Coding System (HCPCS) codes, Uniform Billing (UB) Revenue Codes and ICD-9/ICD-10 procedure codes.

Value sets for the measure are attached in the Excel workbook provided for question S.2b. We include 2016 HEDIS value sets because we used these value sets in measure testing. HEDIS value sets are used because they represent an existing set that states are already familiar with, they are an element of harmonizing with other endorsed measures, and they are updated by the National Committee on Quality Assurance (NCQA). Also, some states may need to include relevant state-specific codes.

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

Adult Medicaid beneficiary discharges from inpatient or residential treatment for SUD with a principal diagnosis of SUD during from January 1 to December 15 of 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).

Population: Medicaid beneficiaries age 18 through 64 as of January 1 of the measurement year.

Benefit: Medical and Behavioral Health Services.

Continuous Enrollment: Date of the inpatient or residential SUD treatment discharge through end of the following month. The enrollment requirement is to ensure that beneficiaries are enrolled for sufficient time to allow for the continuity activities, particularly for a discharge that occurs near the end of a month.

Diagnosis Criteria: Discharges from inpatient or residential treatment with a primary diagnosis of SUD on any claim during the stay. Residential treatment is identified using the value sets in Tabs 1-3 of the attached Excel file. SUD diagnoses are identified using the value sets in Tabs 1-2.

The denominator for this measure is based on discharges, not individuals. If a person has more than one discharge, include all discharges on or between January 1 and December 15 of the measurement year. December 15th is selected to allow sufficient time for continuity activities.

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

Exclude from the denominator for both rates:

- Discharges with hospice services during the measurement year
- Both the initial discharge and the admission/direct transfer discharge if the admission/direct transfer discharge occurs after December 15 of the measurement year.

Discharges followed by admission or direct transfer to inpatient or SUD residential treatment setting within 7 or 14-day continuity of care period. These discharges are excluded from the measure because transfer, hospitalization or admission to residential treatment within 7 or 14 days may prevent a continuity of care visit from taking place. An exception is admission to residential treatment following discharge from inpatient treatment; we do not exclude these admissions, because continuity into residential treatment after inpatient treatment is considered appropriate treatment.

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.)*

Codes reflecting exclusions are attached in S.2b. Residential treatment is identified using the value sets in Tabs 1-3 of the attached Excel file. SUD diagnoses are identified using the value sets in Tabs 1-2.

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.

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.)*

In the steps below we reference the Excel workbook we attached for S.2b. The workbook includes:

- 2016 HEDIS value sets used in measure testing
- 2018 HEDIS value sets used in measure testing for pharmacotherapy and telehealth codes
- Value sets developed during the specification and testing of this measure, and the value sets from NQF #3312 Continuity of Care for Medicaid Beneficiaries After Detoxification (Detox) from Alcohol and/or Drugs and NQF #3400 Use of Pharmacotherapy for Opioid Use Disorder (OUD) that were used in the specification of this measure.

Note - some states may need to also include relevant state-specific codes.

Step 1: Identify denominator

Step 1A. Eligible population: : Identify non-dually enrolled Medicaid beneficiaries age 18 through 64 years with any discharges from inpatient or residential treatment with a principal diagnosis of SUD during January 1 - December 15 of the measurement year. Patients must meet enrollment criteria, defined as Medicaid as the first payer and enrolled in the month of discharge and the following month. Age is calculated as of January 1 of the measurement year.

Throughout Steps 1 and 2, the principal diagnosis of SUD is identified using a principal diagnosis from the 2016 "HEDIS AOD Dependence" value set (Tab 1 in the attached Excel file) or any procedure code from the 2016 "HEDIS AOD Procedures" value set (Tab 2). Secondary diagnosis of SUD is identified using the same value sets.

Step 1B. Flag claims as inpatient or as residential treatment: Among the Medicaid beneficiaries in Step 1A, flag claims as being either in an inpatient or residential setting using all inpatient, outpatient, and ambulatory claims files or tables that contain HCPCS, ICD-9/ICD-10 procedure or diagnosis codes, place of service, or UB revenue codes. Residential treatment is identified using the codes in the SUD Residential Treatment value set (Tab 3). If more than one discharge in a year, treat each discharge as a separate episode, e.g., an inpatient hospital discharge in January and a residential treatment discharge in July counts as two episodes.

Step 1B.1: Consolidate episodes: Multiple inpatient or residential treatment claims that are up to 2 days apart should be combined into a single episode. To facilitate this consolidation, sort the inpatient, outpatient and

ambulatory discharges by Beneficiary ID and service dates to ensure the discharges from these multiple data sources are in chronological order. Use all inpatient and residential treatment claims, regardless of diagnosis, to create episodes.

Step 1C: Assign treatment location to episodes: Use HCPCS, ICD-9/ICD-10 procedure or diagnosis codes, place of service, or UB revenue codes in the SUD Residential Treatment value set (Tab 3) and the SUD diagnosis value sets as noted in Step 1A to assign each episode as inpatient residential treatment, or a mix of both (also indicating the first setting of each episode and the last setting of each episode).

Step 1D: Exclusions: Exclude discharges that meet the exclusion criteria as specified in the "Denominator Exclusion Details" section.

- Exclude discharges for patients who receive hospice services during the measurement year.
- Exclude discharges after December 15 of the measurement year.
- Exclude discharges followed by admission or direct transfer to an inpatient or SUD residential treatment setting within the 7- or 14-day continuity of care period regardless of the principal diagnosis (with exception of admission to residential treatment following discharge from inpatient treatment).
- Exclude episodes that do not include at least one claim with primary diagnosis of SUD.

The denominator for the 7- and 14-day continuity of care rates will differ because of the different exclusions based on transfer or admission to hospital or residential treatment for 7 versus 14 days. For example, a beneficiary admitted to a residential setting on day 10 after discharge will be excluded from the 7-day rate but not from the 14-day rate.

Step 2: Identify numerator

Step 2A: From the denominator, identify discharges from inpatient or residential treatment for SUD with a principal diagnosis of SUD with qualifying continuity of care for SUD (principal or secondary diagnosis) within 7 or 14 days of discharge.

Step 2A.1: Visits: Identify visits meeting continuity of care criteria using outpatient claims files or tables that contain diagnosis, procedure, or revenue codes, procedure code modifiers, or place of service codes. SUD diagnoses can be in any position – primary or secondary – for continuity services. Visits have to occur the day after discharge through day 7 or 14. We identify visits as:

1. Any procedure code or UB revenue code from "HEDIS IET Stand Alone Visits" value set (Tab 4); or
2. Any procedure code from "HEDIS IET Visits Group 1" value set (Tab 5) along with place of service from "HEDIS IET POS Group 1" value set (Tab 6); or
3. Any procedure code from "HEDIS IET Visits Group 2" value set (Tab 7) along with place of service from "HEDIS IET POS Group 2" value set (Tab 8).

The claim must also have procedure code modifier that is missing or a value other than those in the "HEDIS Telehealth Modifier" value set (Tab 9).

Step 2.A.2. Telehealth: Identify visits for telehealth meeting continuity of care criteria using outpatient claims files or tables that contain diagnosis, procedure, or revenue codes, procedure code modifiers, or place of service codes. SUD diagnoses can be in any position – primary or secondary – for continuity services. Telehealth has to occur the day after discharge through day 7 or 14. We identify telehealth as:

1. Any procedure code from the "HEDIS Telephone Visit" value set (Tab 12); or
2. Any procedure code or UB revenue code from "HEDIS IET Stand Alone Visits" value set (Tab 4); or
3. Any procedure code from "HEDIS IET Visits Group 1" value set (Tab 5) along with place of service from "HEDIS IET POS Group 1" value set (Tab 6); or
4. Any procedure code from "HEDIS IET Visits Group 2" value set (Tab 7) along with place of service from "HEDIS IET POS Group 2" value set (Tab 8).

Claims identified using logic in #2-4 must also have procedure code modifier from the "HEDIS Telehealth Modifier" value set (Tab 9).

Step 2A.3: Identify pharmacotherapy events: Indications of pharmacotherapy can occur in outpatient or pharmacy files or tables that contain procedure codes or NDCs. Pharmacotherapy events could be provided on the same day as the discharge through day 7 or 14. Pharmacotherapy continuity claims are identified as follows:

1. In OT file, a) any procedure code from "HEDIS Medication Assisted Treatment" value set (Tab 10); or b) any HCPCS procedure code from "MAT Additional Codes" value set (Tab 11) (developed as part of testing for NQF 3312); or c) any state-specific procedure code from "MAT Additional Codes" value set (Tab 11) for the two states listed in the value set (these codes were identified through consultation for these states).
2. In RX file, any NDC from "AOD Pharmacotherapy" value set (Tab 13). This value set contains NDCs identified as part of testing for NQF 3312 and 3400.

Step 3: Calculate rate

Step 3A: Calculate the overall 7- or 14-day continuity of care rate by dividing the number of discharges with evidence of a qualifying continuity of care visit or pharmacotherapy event (Step 2A) by the denominator (after exclusions) (Step 1D). Calculate the rates separately for each continuity of care time period.

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.

Not applicable. The measure uses administrative data and is not based on a sample.

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.

Claims

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.

Medicaid Alpha-MAX 2014 data: eligible (EL), inpatient (IP), other services (OT), long-term care (LT) and drug (RX) files. The other services (OT) file contains facility and individual provider services data. Most notably, it may contain both residential and other stayover service claims data as claims are assigned to MAX claims file types based upon the category of service provided.

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)*

Population : Regional and State

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

Emergency Department and Services, Home Care, Inpatient/Hospital, Outpatient Services

If other:

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

Not applicable.

2. Validity – See attached Measure Testing Submission Form

SUD-18_Measure_Testing_Attachment_FINAL_SUD_team_07.25.18-636686228615027071.docx

2.1 For maintenance of endorsement

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 For maintenance of endorsement

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 For maintenance of endorsement

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.

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

Measure Number (*if previously endorsed*)**:** 3453
Measure Title: Continuity of care after inpatient or residential treatment for substance use disorder (SUD)
Date of Submission: 7/31/2018

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 ALL 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. If there are differences by aspect of testing,(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 all 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:	☐ other:

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*).

We used the following Medicaid Analytic eXtract (MAX) Medicaid files to identify adult Medicaid beneficiaries with discharges from inpatient or residential SUD treatment (denominator) and whether there was continuity of care treatment services after discharge (numerator):

- Person Summary (PS): Person-level file, including Medicaid eligibility and demographic information
- Inpatient (IP): Claims-level file, including information on inpatient hospital stays
- Long-Term Care (LT): Claims-level file, including information on long-term care institutional stays (nursing facilities, intermediate care facilities for individuals with intellectual disabilities, psychiatric hospitals, etc.)
- Other Therapy (OT): Claims-level file, including information on use of "other" services, such as home- and community-based service use
- Prescription Drug (RX): Information on drugs and other services provided by a pharmacy

1.3. What are the dates of the data used in testing? January-December 2014

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: State	☒ other: State

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*)

We included 13 states in measure testing. State names are redacted.

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*)

Percentage of discharges from an inpatient or residential treatment for substance use disorder (SUD) for Medicaid beneficiaries, ages 18 to 64, which was followed by a treatment service for SUD. SUD treatment includes having an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth encounter, or filling a prescription or being administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

Our testing sample population included all Medicaid beneficiaries (not dual-eligible Medicaid and Medicare) who were between the age 18 and 64, and did not receive hospice services during the measurement year, and had at least one inpatient or residential discharge with a claim that had a primary SUD diagnosis in calendar

year 2014 in 13 states (n=65,820). Table 1 describes the beneficiaries included in the initial analytic sample. State names are redacted.

Table 1. Analytic Sample Selection (1/1/2014 to 12/30/2014)

State	Total enrollment	Total enrollment (age 18 and over)	Total enrollment (ages 18 to 64) who are NOT dually eligible for Medicaid and Medicare	Total enrollment (ages 18 to 64) who are NOT dually eligible for Medicaid and Medicare AND did not receive hospice services	Medicaid beneficiaries with at least one inpatient or residential discharge with a claim that had a primary SUD diagnosis
	(N)	(N)	(N)	(N)	(N)
Total	21,060,599	11,797,744	8,674,514	8,619,565	65,802
State A	846,076	552,739	376,252	375,874	2,782
State B	2,016,557	808,995	480,540	478,774	1,077
State C	596,885	344,014	259,435	259,076	2,447
State D	1,363,680	610,160	393,786	392,998	1,812
State E	2,475,476	1,404,388	1,086,448	1,084,912	6,988
State F	774,696	358,314	190,243	189,540	837
State G	1,117,867	549,213	337,607	336,565	899
State H	6,296,755	4,017,865	3,355,738	3,310,897	27,510
State I	1,115,164	717,397	593,201	592,930	5,892
State J	2,337,475	1,295,863	807,363	805,676	11,434
State K	1,355,998	650,183	424,925	423,805	1,511
State L	184,781	120,563	88,534	88,454	1,758
State M	579,189	368,050	280,442	280,064	855

Note: Based on analysis of 2014 MAX PS, IP, LT, and OT files

The final analytic sample (Table 2) is limited to discharges with Medicaid eligibility in the month of the discharge and the following month, as well as discharges that did not include a direct transfer. One exception is discharges resulting from a direct transfer from an inpatient to a residential setting and did not include an admission within 7 or 14 days. The final sample for 7-day continuity is 59,821 beneficiaries and for 14-day continuity is 58,900 beneficiaries. Table 2 shows that approximately two-thirds of the beneficiaries with at least one inpatient or residential discharge with a claim that had a primary SUD diagnosis during the measurement year in the analytic sample were eligible for Medicaid under the "adult" eligibility category. Slightly more than half of the beneficiaries with at least one discharge in the analytic sample were ages 25 to 44 whereas 37 percent of beneficiaries with at least one discharge in the analytic sample were ages 45-64 years. Sixty-two percent of beneficiaries with at least one discharge in the analytic sample were male. White beneficiaries accounted for a little over a half of beneficiaries with at least one discharge in the analytic sample, followed by Black and Hispanic (23 and 11 percent, respectively).

Table 2. Analytic Sample Demographic Information (7-Day Measure)

	7-Day Rate		14-Day Rate	
Beneficiary Characteristics	**Number of beneficiaries with at least one inpatient or residential discharge with a claim that had a primary SUD diagnosis**	**Percent of beneficiaries with at least one inpatient or residential discharge with a claim that had a primary SUD diagnosis**	**Number of beneficiaries with at least one inpatient or residential discharge with a claim that had a primary SUD diagnosis**	**Percent of beneficiaries with at least one inpatient or residential discharge with a claim that had a primary SUD diagnosis**
	(N)	**(%)**	**(N)**	**(%)**
TOTAL	59,821	100.00%	58,900	100.00%
Medicaid beneficiary category				
Aged	1,161	1.94%	1,157	1.96%
Blind-disabled	16,616	27.78%	16,222	27.54%
Adult, non-disabled	40,927	68.42%	40,417	68.62%
Child	1,117	1.87%	1,104	1.87%
Age				
18-24	6,814	11.39%	6,737	11.44%
25-44	30,716	51.35%	30,318	51.47%
45-64	22,291	37.26%	21,845	37.09%
Gender				
Male	36,984	61.82%	36,389	61.78%
Female	22,837	38.18%	22,511	38.22%
Unknown	0	0.00%	0	0.00%
Race/ethnicity				
White	34,492	57.66%	34,002	57.73%
Black	13,844	23.14%	13,606	23.10%
American Indian/Alaskan Native	537	0.90%	530	0.90%
Asian	730	1.22%	717	1.22%
Hispanic/Latino	6,774	11.32%	6,641	11.28%
Native Hawaiian/Pacific Islander	73	0.12%	72	0.12%
Other Race/Ethnicity	161	0.27%	156	0.26%
Unknown Race/Ethnicity	3,210	5.37%	3,176	5.39%
Rural/Urban				
Rural	7,713	12.89%	7,626	12.95%
Urban	52,008	86.94%	51,178	86.89%
Unknown	100	0.17%	96	0.16%

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 data sample used for different aspects of 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.

As described in section 1.6, we collected information on the following variables using data extracted from Medicaid Analytic eXtract (MAX) 2014 files: Medicaid eligibility category, age, gender, race/ethnicity, and urban/rural during the year.

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*)

Signal-to-noise reliability. The signal-to-noise ratio (SNR) statistic, R (ranging from 0 to 1), summarizes the proportion of the variation between state scores that is due to real differences in underlying entity characteristics (such as differences in population demographics or medical care) as opposed to background-level or random variation (for example, due to measurement or sampling error). As such, it is measure of precision for a given measured entity. If R = 0, there is no true variation on the measure across entities, and all observed variation is due to sampling variation. In this case, the measure is not useful for distinguishing between entities with respect to healthcare quality performance. Conversely, if R = 1, all entity scores are free of sampling error, and all variation represents real differences between entities in the measure result.

We estimated SNR reliability for the SUD-18 measure by first estimating the "noise" (within-state variability), adjusted for the number of beneficiaries within that plan, and estimated the "signal" (between-state variability). We computed the SNR statistic, R (Adams, 2009, 2014), as the ratio of the signal variance (which is common across all entities) to the sum of the signal variance and the noise variance (which varies by entity):

$$R = \frac{\sigma^2_{between}}{\sigma^2_{between} + \sigma^2_{within}}$$

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*)

Signal-to-noise reliability. The SUD-18 measure was highly reliable across the 13 states in our sample, with an average state-level reliability score of 0.989 across states with a range from 0.970 to 0.999 (Table IV.1) for 7-day continuity of care. The 14-day measure of continuity of care reliability averaged 0.990 and ranged from 0.975 to 0.999 (Table IV.2). Note that high reliability is not indicative of high quality of health care, but rather indicates that the SUD-18 measure can be used to meaningfully distinguish the measure's performance across states.

Table IV.1. Signal-to-noise ratio SUD-18 7-day rate, by state

Redacted State Name	N discharges	N recipients of continuity of care	SUD-18 measure rate	Signal-to-noise reliability
Total	84375	15508	18.4	0.989
State A	4186	863	20.6	0.994
State B	1178	130	11.0	0.988
State C	2919	869	29.8	0.990
State D	2036	181	8.9	0.994
State E	8220	1114	13.6	0.998
State F	1077	235	21.8	0.978
State G	1073	188	17.5	0.981
State H	39123	5932	15.2	0.999
State I	7516	1746	23.2	0.997
State J	12183	2600	21.3	0.998
State K	1594	407	25.5	0.983
State L	2276	934	41.0	0.985
State M	994	309	31.1	0.970

Source: Based on analysis of 2014 MAX PS, IP, LT, OT, and RX files.
Notes:
The signal-to-noise coefficients for State H were truncated to 0.999 rather than rounded to 1.000 to reflect the uncertainty in the estimates.

Total SUD-18 measure rate is a weighted average of state rates.

Table IV.2. Signal-to-noise ratio SUD-18 14-day, by state

Redacted State Name	N discharges	N recipients of continuity of care	SUD-18 measure rate	Signal-to-noise reliability
Total	77635	18745	24.1	0.990
State A	3423	1014	29.6	0.994
State B	1117	169	15.1	0.988
State C	2783	1074	38.6	0.991
State D	1937	255	13.2	0.994
State E	7630	1519	19.9	0.998
State F	1016	265	26.1	0.981
State G	988	234	23.7	0.981
State H	35610	6712	18.8	0.999
State I	7031	2284	32.5	0.997
State J	11451	3283	28.7	0.998
State K	1532	485	31.7	0.985
State L	2192	1121	51.1	0.988
State M	925	330	35.7	0.975

Source: Based on analysis of 2014 MAX PS, IP, LT, OT, and RX files.
Notes:

The signal-to-noise coefficients for State H were truncated to 0.999 rather than rounded to 1.000 to reflect the uncertainty in the estimates.
Total SUD-18 measure rate is a weighted average of state rates

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?*)

We used a signal-to-noise ratio (SNR) to estimate the reliability of SUD-18. We found that the measure is highly precise, with all states achieving a SNR reliability estimate >= 0.90. The high SNR is an indication that the SUD-18 measure can discern the performance between states within high precision.

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)*

Face Validity. To evaluate face validity, we surveyed a multi-stakeholder technical expert panel (TEP) that was convened to provide input and guidance on measure development activities under CMS contract HHSM-500-2013-13011I, Task Order # HHSM-500-T0004. The TEP includes 19 individuals representing consumers, state officials, health plans, provider organizations, researchers, and federal government agencies. Twelve individuals responded to the survey. We asked the TEP to rate their agreement that performance scores on the measure "Continuity of care after inpatient or residential treatment for substance use disorder" can be used to distinguish good from poor quality of care. TEP members rated their agreement using a 4-point scale that ranged from strongly disagree to strongly agree.

Convergent Validity. To investigate SUD-18's convergent validity, we compared state-level performance of SUD-18 to state-level performance on two other measures of similar concepts.

- **NQF 3312**: Percentage of discharges from a detoxification episode for adult Medicaid Beneficiaries, age 18 to 64, that were followed by a treatment service for substance use disorder (including the prescription or receipt of a medication to treat a substance use disorder (pharmacotherapy) within 7 or 14 days after discharge. We used NQF 3312 state performance rates that we calculated in the base year of the project.
- **NQF 0576**: The percentage of discharges for members 6 years of age and older who were hospitalized for treatment of selected mental illness diagnoses and who had an outpatient visit, an intensive outpatient encounter or partial hospitalization with a mental health practitioner within 7 days after discharge. The NQF 0576 data are from the CMS Medicaid/CHIP Health Care Quality Measures dataset, which is comprised of performance rates on frequently reported health care quality measures in the Medicaid/CHIP Child and Adult Core Set (Centers for Medicare & Medicaid Services, 2017)

SUD-18 is aligned with NQF 3312 (Continuity of care after detoxification from alcohol or other drugs, 7 and 14 days), with the exception that SUD-18 includes telehealth as a continuity service while NQF 3312

does not. It is similar to NQF 0576 but departs on specific types of continuity services, and the continuity time period. Whereas SUD-18 includes primary and secondary SUD diagnoses, pharmacotherapy, and telehealth as qualifying continuity services, NQF 0576 does not. Due to TEP and public comments, SUD-18 measures 7 and 14 day continuity of care, while NQF 0576 include 7 and 30 day measures. In addition, NQF 0576 includes beneficiaries for 6 years of age and older while SUD-18 includes discharges for beneficiaries age 18-64.

We assessed the convergent validity of the SUD-18 measure by calculating its Spearman rank correlation with the two external measures. The Spearman rank correlation ranges from −1 to 1, with positive value indicating a positive relation between the two measures and negative value showing an opposite direction of the two. Moreover, large magnitude (regardless of the sign) of the correlation value demonstrates a strong association between the two measures, whereas a correlation value close to zero implies a weak association.

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

Face validity. Twelve out of 12 respondents agreed or strongly agreed that performance scores on the measure "Continuity of care after inpatient or residential treatment for substance use disorder" can be used to distinguish good from poor quality of care. Two TEP members noted their strong assessment only applied to the 7-day and not the 14-day continuity rate.

Convergent validity. The following table lists the performance rates of the three measures for each state. Note that we have 13 states from our analytic data with SUD-18 measure rates. Of these, 12 also appear in the core set database for the NQF 0576 measure and 11 appear in the database for the NQF 3312 measure. Hence, our analysis below focuses on these states.

Table IV.3. Performance rates for SUD-18, NQF 3312, and NQF 0576, by state

Redacted State Name	7-day Rate			14-day Rate	
	SUD-18	NQF 3312	NQF 0576	SUD 18	NQF 3312
State A	20.6	20.3	38.2	29.6	29.7
State B	11.0	11.9	38.0	15.1	19.3
State C	29.8	18.7	38.0	38.6	26.9
State E	13.6	59.2	NR	19.9	64.0
State F	21.8	13.5	25.5	26.1	19.9
State G	17.5	26.1	32.6	23.7	33.0
State H	15.2	14.7	35.2	18.8	22.6
State J	21.3	49.3	43.6	28.7	55.3
State K	25.5	26.4	54.7	31.7	33.0
State L	41.0	79.1	60.9	51.1	81.3
State M	31.1	27.0	6.7	35.7	35.8
State D	8.9	NR	17.8	13.2	NR
State I	23.2	NR	64.4	32.5	NR

Source: Based on analysis of 2014 MAX PS, IP, LT, OT, and RX files. Note: NR = Not reported

The correlation between state-level SUD-18 is strongest with **NQF 3312** (14-day continuity of care). The SUD-18: NQF 3312 measures have a correlation of 0.46 (95 percent confidence interval: -0.19, 0.83), which indicates the measures have a moderate to strong positive correlation (Evans, 1996) (see Table IV.4 below) . Specifically, we find states with low or high **NQF 3312** 14-day rates tend to have high SUD-18 14-day rates as well. Exceptions are states E and L. States E and L have fairly normal rates for SUD-18 14-day, but appear to be outliers for **NQF 3312** 14-day.

The association between the SUD-18 7-day measure and the NQF 3312 7-day measure also indicates a clear positive trend. The correlation coefficient is 0.40 with a confidence interval of (-0.26, 0.81).

While not as strong of an association, the association between the SUD-18 7-day measure and the NQF 0576 measure also suggests a positive trend, with a correlation coefficient of 0.34 (95 percent confidence interval: -0.29, 0.76). The weaker correlation between SUD-18 and NQF 0576 could be due, in part, to age differences in the sample. SUD-18 includes adults 18 to 64 years; NQF 0576 includes adults 21 to 64 in ten states and "other ages" in the remaining three states, which could include beneficiaries 18-21 or over age 65. In addition, for SUD 18 some beneficiaries were discharged from specialty residential SUD settings while beneficiaries in NQF 0576 may have been discharged from specialty psychiatric hospitals. In some states, these may be distinct service systems and the focus on follow-up after discharge may differ between these systems.

Note, the confidence intervals on all three correlations are large, and have a negative lower bound due to the small sample size (n = 11 between SUD-18 and **NQF 3312**, and n = 12 between SUD-18 and SUD 0576).

Table IV.4. Spearman rank correlation coefficients comparing SUD-18 with NQF 3312 and NQF 0576

SUD-18 measure	External measure	Spearman rank correlation	95% CI
SUD-18 7-day	NQF 3312 7-day	0.40	(-0.26 ,0.81)
SUD-18 7-day	NQF 0576 7-day	0.34	(-0.29 ,0.76)
SUD-18 14-day	NQF 3312 14 day	0.46	(-0.19 ,0.83)

Source: Based on analysis of 2014 MAX PS, IP, LT, OT, and RX files.

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?*)

The face validity results suggest SUD-18 is a valid measure of continuity of care after inpatient or residential treatment for SUD. Overall, the consistent directionality of the correlation between SUD-18 and the external variables and the strength of the correlations suggest SUD-18 has moderate convergent validity. Together, the results suggest SUD-18 has moderate validity.

2b2. EXCLUSIONS ANALYSIS

☐ **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*)

We did not test differences in performance with the exclusion of beneficiaries who entered hospice since the number was negligible (0.6 percent of all beneficiaries who are not dually eligible). Furthermore, hospice care is for a different population than those in SUD treatment. Moreover, we could not test the impact of excluding discharges followed by transfer or admission to any inpatient or residential setting in 7 or 14 days because the measure cannot be calculated without these exclusions. That is, beneficiaries transferred or admitted during the 7 or 14 day window are not in the community to allow them the opportunity for a continuity service to take place.

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*)

Not applicable.

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*)

As stated above, the measure cannot be calculated without the exclusions due to transfer or admission within 7 or 14 days.

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.

Not applicable - Not an intermediate or health outcome, or PRO-PM, or resource use measure.

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.

Not applicable

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.

Not applicable

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?

Not applicable

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)**

Not applicable

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

Not applicable

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.**

Not applicable

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

Not applicable

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:** Not applicable

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*): Not applicable

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves: Not applicable

2b3.9. Results of Risk Stratification Analysis: Not applicable

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*). Not applicable

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*). Not applicable

---

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)*

The NQF criteria for the measure's importance include an evaluation of meaningful differences in SUD-18 rates between subgroups of key patient characteristics. To inform the measure's importance, we examined the magnitude of differences across groups by Medicaid beneficiary categories, age, gender, race/ethnicity, and urban/rural residence.

To assess whether SUD-18 performance differs meaningfully among states included in testing, we calculated the 95 percent confidence interval of the SUD-18 measure rate for each state, where the SUD-18 rate for each state can be viewed as a proportion. We then compared each state's confidence interval to the overall measure rate (that is, across all states). States' measure rates that are significantly higher than the overall rate indicate less than optimal performance, which suggests room for improvement. We conducted these analyses for each of the populations. We also analyzed the magnitude of differences in measure performance by population subgroups of interest (age group, sex, race/ethnicity, disability status, and payer).

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*)

We found that the SUD-18 measure rates across the 13 states cover a wide range with meaningful variation. The measure rate ranges from 8.9-percent to 41.0-percent and 24.2-percent to 51.1-percent continuing care, for 7-day and 14-day follow-ups, respectively. States L, M, and C had the highest 7-day and 14-day continuity of care. Five states show significantly lower measure rates than the overall SUD-18 measure rate, indicating room for improvement (see Table III.1, Figure III.1 and Figure III.2).

Table III.1. SUD-18 measure rates, by state

Redacted State Name	Continuity of Care	N discharges	N recipients of continuity of care	SUD-18 rate	95% CI
Total	7-day	84375	15508	18.4	
	14-day	77635	18745	24.1	
State A	7-day	4186	863	20.6	(19.39, 21.84)
	14-day	3423	1014	29.6	(28.09, 31.15)
State B	7-day	1178	130	11.0	(9.25, 12.82)
	14-day	1117	169	15.1	(13.03, 17.23)
State C	7-day	2919	869	29.8	(28.11, 31.43)
	14-day	2783	1074	38.6	(36.78, 40.40)
State D	7-day	2036	181	8.9	(7.65, 10.13)
	14-day	1937	255	13.2	(11.66, 14.67)
State E	7-day	8220	1114	13.6	(12.81, 14.29)
	14-day	7630	1519	19.9	(19.01, 20.80)
State F	7-day	1077	235	21.8	(19.35, 24.29)
	14-day	1016	265	26.1	(23.38, 28.78)
State G	7-day	1073	188	17.5	(15.25, 19.80)
	14-day	988	234	23.7	(21.03, 26.34)
State H	7-day	39123	5932	15.2	(14.81, 15.52)
	14-day	35610	6712	18.8	(18.44, 19.25)
State I	7-day	7516	1746	23.2	(22.28, 24.19)
	14-day	7031	2284	32.5	(31.39, 33.58)
State J	7-day	12183	2600	21.3	(20.61, 22.07)
	14-day	11451	3283	28.7	(27.84, 29.50)
State K	7-day	1594	407	25.5	(23.39, 27.67)
	14-day	1532	485	31.7	(29.33, 33.99)
State L	7-day	2276	934	41.0	(39.02, 43.06)
	14-day	2192	1121	51.1	(49.05, 53.23)
State M	7-day	994	309	31.1	(28.21, 33.96)
	14-day	925	330	35.7	(32.59, 38.76)

Figure III.1. SUD-18 7-day Measure rate exhibits significant and clinically meaningful differences between states

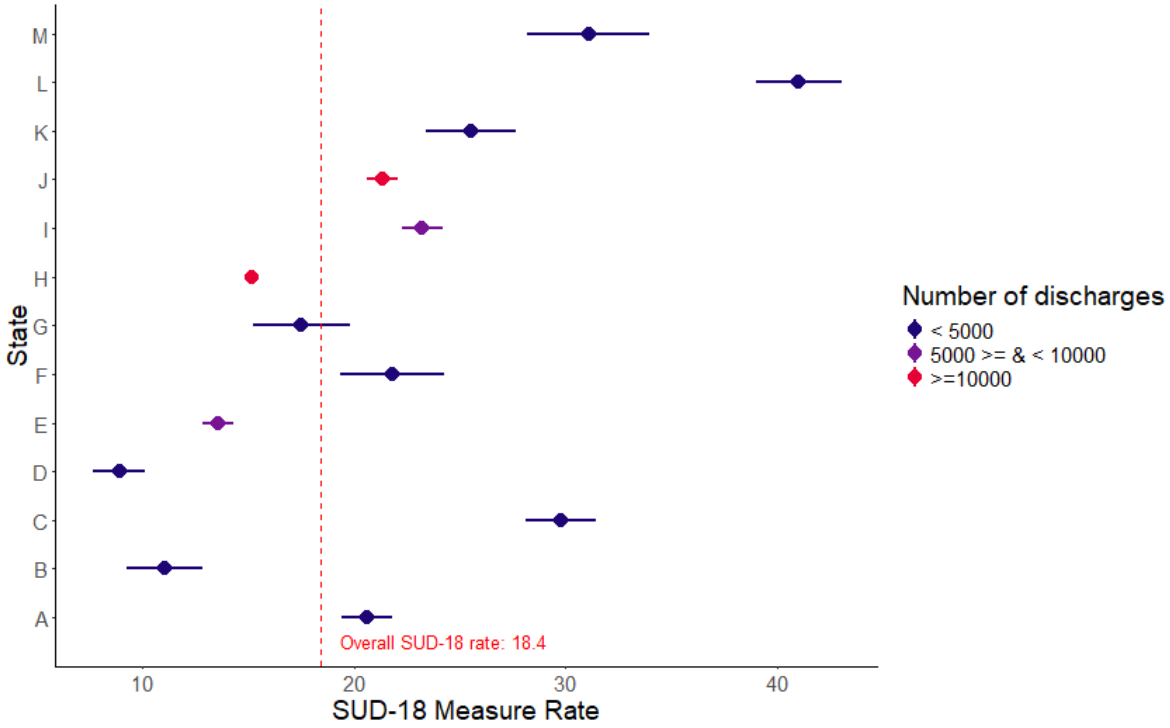
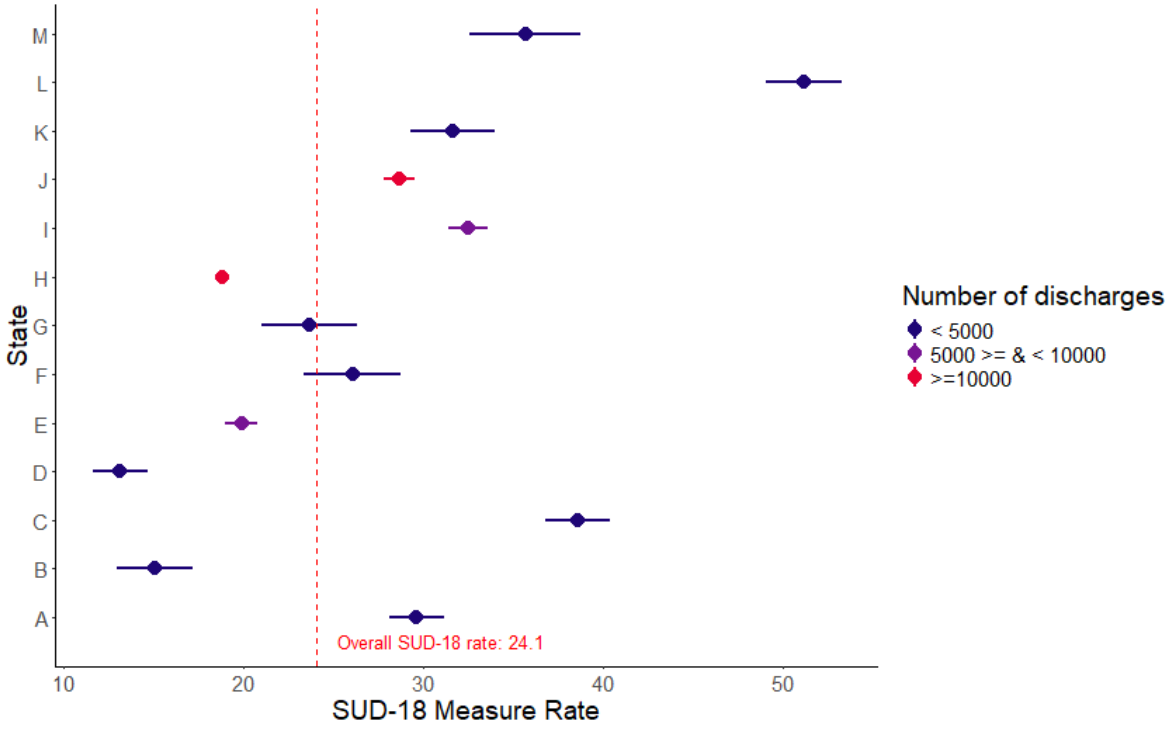


Figure III.2. SUD-18 14-day Measure rate exhibits significant and clinically meaningful differences between states



Meaningful (differences across subpopulations)

We conducted tests to determine whether there are meaningful differences across Medicaid beneficiary categories, age, gender, race/ethnicity, and urban/rural residence. By conducting a Chi-squared test on each of these patient characteristics, we found significant variation in continuity of care rates between subgroups of each patient characteristic. Taken as a whole, these analyses demonstrate that meaningful variation exists in the underlying measure rates, both for the total and across states. As a post-hoc analysis, we also conduct a proportion test between each pair groups (e.g. age 18-24 vs 45-64) within each category to investigate the differences between each pair. We highlight the findings for each category in the following:

In interpreting the results, we note that, some of the subgroup categories have a high SUD-18 rate but represent a low percentage of the population. For example, the SUD-18 rate for aged Medicaid beneficiaries is 38.8 percent, compared to an overall weighed mean of 18.4 for the Medicaid beneficiary group. However, aged beneficiaries make up only 1.9 percent of the total discharges.

Age. The sample mean indicates that beneficiaries age 25-44 were more likely to have continuity of care visits than older beneficiaries. There was a statistically significant difference between age groups 45-64 and 18-24, and 45-64 and 25-44. However, there is no significant difference between age groups 25-44 and 18-24.

Sex. Females appear to attend follow-ups at higher rates than males, and the difference is statistically significant.

Race/ethnicity. Whites and American Indian have the highest rates of continuing care, 28.4 and 27.8 percent (14-day follow-up), respectively. Blacks and Asians have the lowest rates of continuing care, 15.3 and 11.9 percent (14-day follow-up), respectively. There are significant differences in continuity of care rates between White and Asian and White and Black. Note that we found a relatively high proportion of claims with missing race (missing as high as 37 percent in one state), therefore, findings should be interpreted with caution.

Medicaid beneficiary categories. There are significant differences between many of the categories. Aged had the highest mean continuity of care rate (39-percent), but makes up a very small percent of the population.

Other categories had much lower continuity of care rates (15-19 percent). There are significant differences between aged and all other categories and significant differences between adult, non-disabled and blind-disabled.

Rural/urban. There are large, significant differences between discharges living in urban/rural environments and continuing care. For both 7-day and 14-day rates, discharges living in rural environments continued care at a much higher rate than those living in urban environments (32.4 percent compared to 23.0 percent, respectively for the 14-day rate).

Table III.5. Distribution of beneficiaries with 7- and 14-day continuity of care, by beneficiary characteristics

Category	Subgroup	7-Day Rate		14-Day Rate	
		N discharges	Weighted mean	N discharges	Weighted mean
Age***	Total	84375	18.4	77635	24.1
	18-24	8692 (10.3%)	19.8	8198 (10.6%)	25.8
	25-44	42322 (50.2%)	20.5	39163 (50.4%)	26.7
	45-64	33361 (39.5%)	15.3	30274 (39.0%)	20.3
Sex***	Total	84375	18.4	77635	24.1
	Female	30014 (35.6%)	21.9	28110 (36.2%)	28.7
	Male	54361 (64.4%)	16.5	49525 (63.8%)	21.6
Race/Ethnicity***	Total	84375	18.1	77635	24.0
	American Indian or Alaskan Native	731 (0.9%)	19.2	681 (0.9%)	27.8
	Asian	1202 (1.4%)	9.1	1066 (1.4%)	11.9
	Black	20754 (24.6%)	11.3	18762 (24.2%)	15.3
	Hispanic/Latino	10,536 (12.5%)	20.6	9,495 (12.2%)	24.9
	Native Hawaiian or Pacific Islander	114 (0.1%)	17.5	106 (0.1%)	24.5
	Other Race	199 (0.2%)	22.1	186 (0.2%)	28.0
	Unknown Race	4116 (4.9%)	14.9	3898 (5.0%)	20.1
	White	46723 (55.4%)	21.6	43441 (56.0%)	28.4
Medicaid beneficiary category***	Total	84375	18.4	77635	24.1
	Adult, Non-disabled	57496 (68.1%)	18.8	52974 (68.2%)	24.5
	Aged	1615 (1.9%)	38.8	1545 (2.0%)	49.0
	Blind-disabled	23851 (28.3%)	16.2	21788 (28.1%)	21.8
	Child	1413 (1.7%)	14.8	1328 (1.7%)	19.1
Urbanicity***	Total	84375	18.4	77635	24.1
	Rural	9572 (11.3%)	24.5	9174 (11.8%)	32.4
	Urban	74681 (88.5%)	17.6	68345 (88.0%)	23.0
	Urban Rural Unknown	122 (0.1%)	32.8	116 (0.1%)	40.5

***< 0.01

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?*)

Taken as a whole, these analyses demonstrate that meaningful variation exists in the underlying measure rates, both for the total and across states, and by beneficiary subgroups. The measure results suggest variation in performance and room for improvement in continuity of care rates. When looking at state-specific SUD-18 measure rates, 5 of the 13 states, or 38 percent, exhibit significantly lower measure rates than average.

Overall, wide variation across states and generally low performance indicates room for improvement in the care provided for beneficiaries with SUD with respect to follow-up care.

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.***

Not applicable.

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

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*). Not applicable

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*). Not applicable

---

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*)

We assessed missing data using MAX Validation Reports.[1] These tables are used to evaluate the quality and completeness of MAX data generally. The eligibility and utilization data elements used in the calculation of SUD-18 (dates and place of service, diagnosis and procedure codes, and revenue center codes) – are generally required for either the payment of claims or for inclusion in MAX files.[2] There is therefore little missing data for these data elements in their respective claims files, and results are therefore unlikely to be biased due to missingness. The one variable in the meaningful differences analyses that has substantial missing data is race/ethnicity; the completeness of this variable is not required for claims payment or inclusion in the MAX files.

---

[1] MAX validation tables and further information can be found here: https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAX-Validation-Reports.html.

[2] This varies by claim type. We expect procedure code to be fully populated on OT claims because they are at the service level and a procedure code is required for payment. For IP claims, however, procedure code is optional but diagnosis code is required.

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; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Tables V.6 and V.7 present information on missingness for the data elements used in measure calculation from the MAX validation tables. We find a relatively high proportion of claims with missing race; for example, the percent with missing race is as high as 37 percent in one state. Any calculations for race/ethnicity subgroups do not include beneficiaries for which race/ethnicity is missing and therefore may be biased.

Table V.6. Quality and completeness measures for eligibility and IP claims data from the MAX Validation Reports

State	PS All Records: % with No Claims (Recipient Indicator = 0)	PS All Records: % with Claims and Missing Medicaid Eligibility (Excludes S-CHIP Only)	PS Enrolled: % Unknown Race	PS Enrolled: % with Gender Code M or F	IP Stays: % Missing Eligibility and > $0 Paid (Excludes S-CHIP Only)	IP FFS Non-crossover: % Stays with Primary Diagnosis Code	IP FFS Non-crossover: % Stays with a Procedure Code	IP Encounter: % Stays with Primary Diagnosis Code	IP Encounter: % Stays with a Procedure Code
State A	13.19	0.27	0.00	100.00	0.22	100.00	58.37	NA	NA
State B	14.58	0.85	9.99	100.00	0.07	100.00	60.74	100.00	73.94
State C	4.15	0.08	37.21	100.00	0.04	100.00	62.15	100.00	0.02
State D	13.81	4.40	9.74	100.00	1.62	100.00	63.06	100.00	57.20
State E	6.23	1.50	15.24	100.00	0.94	100.00	65.10	100.00	62.75
State F	13.13	0.35	6.58	100.00	0.28	100.00	31.95	100.00	46.26
State G	18.09	0.23	6.40	100.00	0.21	100.00	44.13	100.00	61.95
State H	12.32	0.07	7.68	99.02	0.23	100.00	74.83	100.00	77.49
State I	15.86	0.35	19.68	100.00	0.02	100.00	58.30	100.00	55.68
State J	6.14	3.77	12.66	100.00	0.94	100.00	66.87	100.00	66.28
State K	7.34	0.56	16.89	100.00	0.00	NA	NA	98.10	60.61
State L	10.40	0.22	23.58	100.00	0.41	100.00	58.94	NA	NA
State M	6.32	0.09	0.16	100.00	0.05	100.00	60.74	NA	NA

Source: MAX validation tables. Available at the following URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAX-Validation-Reports.html. We use the most recently available tables. For three states these data are from 2013, for all others it is from 2014.

NA = Not Applicable.

Note: Crossover claims are for beneficiaries eligible for both Medicaid and Medicare. Because we exclude these dually eligible beneficiaries from the analysis, we only examined missingness for non-crossover claims.

Table V.7. Quality and completeness for LT and OT claims data from the MAX Validation Reports

State	LT All Claims: % Missing Eligibility and > $0 Paid (Excludes S-CHIP Only)	LT FFS Non-crossover: % Claims with Primary Diagnosis Code	LT Encounter: % Claims with Primary Diagnosis Code	OT All Claims: % Missing Eligibility and > $0 Paid (Excludes S-CHIP Only)	OT FFS Non-crossover: % Claims with Place of Service	OT FFS Non-crossover: % Claims with Primary Diagnosis Code	OT FFS Non-crossover: % Claims with Procedure Code	OT Encounter: % Claims with Primary Diagnosis Code	OT Encounter: % Claims with Procedure Code
State A	0.07	100.00	NA	0.18	92.29	88.79	91.33	NA	NA
State B	0.01	100.00	100.00	0.16	88.31	96.14	96.58	78.14	99.89
State C	0.02	99.96	NA	0.00	90.96	90.79	100.00	100.00	100.00
State D	0.12	86.68	100.00	3.11	88.85	93.65	99.31	89.53	99.89
State E	0.46	100.00	100.00	0.10	99.89	76.75	99.64	96.81	97.71
State F	0.02	100.00	100.00	0.10	80.27	83.23	99.55	82.63	96.13
State G	0.02	100.00	NA	0.07	92.59	97.71	100.00	83.16	99.14
State H	0.21	100.00	99.97	0.00	87.64	97.37	99.24	91.55	99.05
State I	0.01	100.00	100.00	0.04	97.07	73.21	99.66	88.87	98.24
State J	0.16	100.00	100.00	0.31	70.61	97.55	100.00	79.30	99.03
State K	0.00	100.00	99.99	0.03	99.96	60.65	100.00	99.93	99.88
State L	0.29	100.00	NA	0.04	93.47	97.47	92.74	NA	NA
State M	0.02	100.00	NA	0.01	98.06	96.93	98.72	NA	NA

Source: MAX validation tables. Available at the following URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAX-Validation-Reports.html. We use the most recently available tables. For three states these data are from 2013, for all others it is from 2014.

NA = Not Applicable.

Note: Crossover claims are for beneficiaries eligible for both Medicaid and Medicare. Because we exclude these dually eligible beneficiaries from the analysis, we only examined missingness for non-crossover claims.

The MAX validation tables do not contain information on claims or encounters that are not in the MAX files. There are a number of reasons that claims or encounters might not make it into MAX data, including but not limited to: (a) if they were not fully or correctly populated, or (b) if a managed care or behavioral health organizations were unable to submit their encounters in an acceptable format. Unfortunately, we are not able to determine how many claims and encounters may be left out of the MAX data for these reasons. However, we can look at the number and percent of beneficiaries in FFS and managed care without a claim as a proxy for "missing" encounters. We do this for OT claims in particular, as we expect the vast majority of beneficiaries to have at least one OT claim in the measurement year.

Since behavioral health claims in particular can pose a data completeness and/or quality challenge, we investigated whether claims for beneficiaries in BHOs are underreported in states with a large proportion of their beneficiaries enrolled in a BHO. In particular, we examined the percent of beneficiaries who do not have any OT claim in 2014. We determined the vast majority of behavioral health encounters from these states are included in MAX data.

In addition to missingness, it is well known that some states use local revenue center codes. These data will not appear to be missing in the validation tables, but we will not be able to use them with the standard value sets. Like all measures calculated using claims data, states using local codes should include those codes in calculating the measure as applicable.

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; if no empirical analysis, provide rationale for the selected approach for missing data*)

Given the relatively small amount of missing information, we don't believe there is systematic bias. In addition, states implementing the measure will likely have even less missing data than reported here because they will be able to account for their state-specific codes when constructing the measure.

References

Adams, J. L. (2009). The Reliability of Provider Profiling. A Tutorial. http://www.rand.org/pubs/technical_reports/TR653.html

Adams, J. L. (2014). Reliability-Testing Concepts. National Quality Forum presentation. Retrieved from http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=74717

Centers for Medicare & Medicaid Services. (2017). 2017 Core Set of Behavioral Health Measures for Medicaid and CHIP (Behavioral Health Core Set). Retrieved from https://www.medicaid.gov/medicaid/quality-of-care/downloads/2017-bh-core-set.pdf

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing.

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.

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

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**.

ALL data elements are in defined fields in electronic claims

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).

Not applicable.

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.

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)
Regulatory and Accreditation Programs Quality Improvement (Internal to the specific organization)	

4a1.1 For each CURRENT use, checked above (update for <u>maintenance of endorsement</u>), 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; the measure is under initial endorsement review and is not currently used in an accountability program.

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?*)

CMS is considering implementation plans for this measure. There are no identified barriers to implementation in a public reporting or accountability application.

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.*)

CMS is developing measures to improve the quality of care of the following Medicaid populations served by CMS's Innovation Accelerator Program:

• People eligible for both Medicare and Medicaid, or "Dual eligible beneficiaries"

• People receiving long-term services and supports (LTSS) through managed care organizations

• People with substance use disorders; beneficiaries with complex care needs and high costs; beneficiaries with physical and mental health needs; or Medicaid beneficiaries who receive LTSS in the community

This measure is intended for voluntary use by states to monitor and improve the quality of care provided for Medicaid beneficiaries with substance use disorders. States may choose to begin implementing the measures based on their programmatic needs.

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 beneficiaries, who are discharged from inpatient or residential treatment for SUD. Currently the overall rate of continuity of care after inpatient and residential treatment is 16.3% and 22.3%, respectively, for the 7-day measure across the 13 states included in testing. The inpatient 7-day rate ranges from 8.9% in State D to 40.9% in State L, and the residential rate ranges from 8.9% in State D to 87.9% in State M, indicating an opportunity for improvement. . The Continuity of care after inpatient or residential treatment for substance use disorder measure may be useful for monitoring the rate of continuing care and encourage states to put interventions in place to increase the rates. This is important because continuity of care (defined in time frames ranging from 7 days to one year post-discharge) has been shown to be related to better outcomes such as reduction in substance use, relapse, readmissions, and criminal justice involvement while also related to improved employment status (Blodgett, Maisel, Fuh, Wilbourne, & Finney, 2014; DeMarce et al., 2008; Mark et al., 2013; McKay, 2005, 2009; McKay & Hiller-Sturmhofel, 2011). Continuity of care has also been found to lower the risk of death (Harris et al., 2015) and to sustain treatment gains that are made (Schaefer et al., 2005).

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.

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

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

This is a new measure that has not yet been implemented. No unexpected benefits were observed during testing.

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.

Yes

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

0004 : Initiation and Engagement of Alcohol and Other Drug Abuse or Dependence Treatment

0576 : Follow-Up After Hospitalization for Mental Illness (FUH)

1937 : Follow-Up After Hospitalization for Schizophrenia (7- and 30-day)

2605 : Follow-Up After Emergency Department Visit for Mental Illness or Alcohol and Other Drug Abuse or Dependence

3312 : Continuity of Care for Medicaid Beneficiaries after Detoxification (Detox) From Alcohol and/or Drugs

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

Not applicable.

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?

No

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

Parts of the specifications for the proposed measure harmonize with some measures but not others. Below we describe similarities and differences between the proposed measure and other measures. The differences do not impose additional data collection burden to states, because the data elements are available in administrative data and are consistent with some measures states are already likely collecting. Numerator: Timing of continuity of care. The proposed measure specifies continuity of care within 7- and 14-days of discharge and is harmonized with NQF 3312, Continuity of care for Medicaid beneficiaries after detoxification (detox) from alcohol and/or drugs, which also focuses on a SUD population. NQF 0576, 1937, and 2605 all specify follow-up within 7 and 30 days. The populations for NQF 0576 and 1937 include patients with mental health related diagnoses rather than focusing on substance use disorders. NQF 2605 has a target mixed population of mental health and SUD patients. In measure testing, stakeholders expressed concern that 30 days is too long for SUD patients to wait for a continuity of care service after discharge from inpatient or residential care. Timelier follow-up with these patients is needed so as not to lose them. NQF 0004 is partially harmonized with the proposed measure in that the initiation visit is specified as within 14 days of the index episode start date (diagnosis). Diagnoses in the continuity of care visit. The proposed measure is harmonized with NQF 3312 and NQF 0004 by allowing SUD to either be the primary or a secondary diagnosis for treatment services that count toward continuity in the numerator. This is to address potential inaccuracies in how SUD diagnoses are coded. For example, some providers may be concerned about the stigma associated with an

SUD diagnosis and therefore code it as a secondary diagnosis. Also, for adults with co-occurring mental health and SUD disorders, the assignment of primary and secondary diagnoses can be challenging and sometimes arbitrary. NQF 2605 does not allow a secondary SUD diagnosis. NQF 0576, NQF 1937, are not clear on whether only a primary diagnosis is allowed in the numerator. Services to include as continuity of care. The proposed measure includes pharmacotherapy and telehealth as services that count as continuity of care. NQF 2605, 0576, and 1937 do not include these services. Adding an SUD medication or telehealth claim as evidence of continuity of care is consistent with recent changes made to the 2018 HEDIS specification of NQF 0004 (National Committee on Quality Assurance, 2018). Practitioners valid for providing follow-up services. The proposed measure and NQF 2605 allow any practitioner to provide follow-up services, because of the expectation that the follow-up services captured in the measure may be provided by primary care clinicians. NQF 0576 and 1937 only allow non-mental health practitioners in specified settings and with specific diagnosis codes. Denominator: Diagnoses in denominator. The denominators for the proposed measure and all the related measures are harmonized in requiring a primary diagnosis for the condition that is the measure's focus. Age. The proposed measure is intended for an adult Medicaid population. Similar to NQF 3312 and NQF 1937, it includes ages 18-64.The proposed measure excludes adults over 64 years, because complete data on services received by dually-eligible (Medicaid and Medicare) adults are not available in Medicaid data. NQF 2605 includes adults age 18 and older. NQF 0576 includes individuals age 6 and older and NQF 0004 includes age 13 and older. In terms of impact on interpretability, the proposed measure would have lower continuity rates than the measures that have a 30-day follow-up time period and higher continuity rates than the measures that only count non-mental health practitioners in certain settings and with certain diagnosis codes.

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.)

Not applicable; there are no competing measures.

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.

No appendix Attachment:

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: Melissa, Azur, mazur@mathematica-mpr.com, 202-250-3518-

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.

The project's Technical Expert Panel provided input on measure selection, feedback on testing results, and an assessment of the face validity of performance scores. The TEP includes the following members:

Consumer Representative 1

- Carol McDaid

- Capitol Decisions, Inc

Consumer Representative 2

- Janice Tufte

- Patient-Centered Outcomes Research Institute (PCORI) ambassador

- PCORI

Consumer Representative 3

- Kayte Thomas

- PCORI ambassador

- PCORI

State Official 1

- Joe Parks

- Missouri HealthNet Division (Medicaid)

State Official 2

- David Mancuso

- Washington State Department of Social and Health Services

State Official 3

- Roxanne Kennedy

- New Jersey Division of Mental Health and Addiction Services

Health Plan Representative 1

- Alonzo White

- Aetna Medicaid

Health Plan Representative 2

- Deb Kilstein

- Association for Community Affiliated Plans

Health Plan Representative 3

- Jim Thatcher

- Massachusetts Behavioral Health Partnership, Beacon Health Options

Provider Organization Representative 1

- Daniel Bruns

- Health Psychology Associates

Provider Organization Representative 2

- Aaron Garman

- Coal Country (ND) Community Health Center (and American Academy of Family Practice Comm. on Quality & Practice)

Provider Organization Representative 3

- Annette DuBard

- Community Care of North Carolina

Subject Matter Expert/Researcher 1

- Andrew Bindman

- University of California San Francisco (incoming AHRQ director)

Subject Matter Expert/Researcher 2

- Mady Chalk

- Treatment Research Institute

Subject Matter Expert/Researcher 3

- Kimberly Hepner

- RAND Corporation

Subject Matter Expert/Researcher 4

- Benjamin Miller

- University of Colorado, School of Public Health

Subject Matter Expert/Researcher 5

- Alex Sox-Harris

- Department of Veterans Affairs

Federal Agency Official 1

- Deb Potter

- Office of the Assistant Secretary for Planning and Evaluation

Federal Agency Official 2

- Laura Jacobus-Kantor

- Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality

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? Specifications for this measure will be reviewed and updated annually.

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

Ad.6 Copyright statement: Limited proprietary coding is contained in the Measure specifications for user

convenience. Users of proprietary code sets should obtain all necessary licenses from the owners of the code sets. Mathematica disclaims all liability for use or accuracy of any CPT or other codes contained in the specifications.

CPT(R) contained in the Measure specifications is copyright 2004-2016 American Medical Association.

ICD-10 copyright 2016 World Health Organization. All Rights Reserved.

The American Hospital Association holds a copyright to the National Uniform Billing Committee (NUBC) codes contained in the measure specifications. The NUBC codes in the specifications are included with the permission of the AHA. The NUBC codes contained in the specifications may be used by health plans and other health care delivery organizations for the purpose of calculating and reporting Measure results or using Measure results for their internal quality improvement purposes. All other uses of the NUBC codes require a license from the AHA. Anyone desiring to use the NUBC codes in a commercial product to generate measure results, or for any other commercial use, must obtain a commercial use license directly from the AHA. To inquire about licensing, contact ub04@healthforum.com.

NDC codes change periodically and should be updated whenever the measure is applied.

This measure contains HEDIS® Value Sets that were developed, are owned by and are included with the permission of the National Committee for Quality Assurance ("NCQA"). Proprietary coding is contained in the HEDIS Value Sets. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. NCQA disclaims all liability for use or accuracy of any coding contained in the HEDIS Value Sets. The HEDIS Value Sets are provided "as is" without warranty of any kind. Users shall not have the right to alter, enhance or otherwise modify the HEDIS Value Sets, and shall not disassemble, recompile or reverse engineer the HEDIS Value Sets. All uses of the HEDIS Value Sets outside the measure must be approved by NCQA and are subject to a license at the discretion of NCQA. ©2016 and 2018 NCQA, all rights reserved."

Ad.7 Disclaimers: This performance measures is not a clinical guideline and does not establish a standard of medical care, and has not been tested for all potential applications. The measures and specifications are provided without warranty.

Ad.8 Additional Information/Comments:

References

Blodgett, J. C., Maisel, N. C., Fuh, I. L., Wilbourne, P. L., & Finney, J. W. (2014). How effective is continuing care for substance use disorders? A meta-analytic review. J Subst Abuse Treat, 46(2), 87-97. doi: 10.1016/j.jsat.2013.08.022

DeMarce, J. M., Lash, S. J., Stephens, R. S., Grambow, S. C., & Burden, J. L. (2008). Promoting continuing care adherence among substance abusers with co-occurring psychiatric disorders following residential treatment. Addict Behav, 33(9), 1104-1112. doi: 10.1016/j.addbeh.2008.02.008

Garnick, D. W., Lee, M. T., Horgan, C. M., Acevedo, A., & Washington Circle Public Sector, Workgroup. (2009). Adapting Washington Circle performance measures for public sector substance abuse treatment systems. J Subst Abuse Treat, 36(3), 265-277. doi: 10.1016/j.jsat.2008.06.008

Harris, A. H., Gupta, S., Bowe, T., Ellerbe, L. S., Phelps, T. E., Rubinsky, A. D., . . . Trafton, J. (2015). Predictive validity of two process-of-care quality measures for residential substance use disorder treatment. Addict Sci Clin Pract, 10, 22. doi: 10.1186/s13722-015-0042-5

Harris, A. H., McKellar, J. D., Moos, R. H., Schaefer, J. A., & Cronkite, R. C. (2006). Predictors of engagement in continuing care following residential substance use disorder treatment. Drug Alcohol Depend, 84(1), 93-101. doi: 10.1016/j.drugalcdep.2005.12.010

Mark, T. L., Tomic, K. S., Kowlessar, N., Chu, B. C., Vandivort-Warren, R., & Smith, S. (2013). Hospital readmission among medicaid patients with an index hospitalization for mental and/or substance use disorder. J Behav Health Serv Res, 40(2), 207-221. doi: 10.1007/s11414-013-9323-5

McKay, J. R. (2005). Is there a case for extended interventions for alcohol and drug use disorders? Addiction, 100(11), 1594-1610. doi: 10.1111/j.1360-0443.2005.01208.x

McKay, J. R. (2009). Continuing care research: what we have learned and where we are going. J Subst Abuse Treat, 36(2), 131-145. doi: 10.1016/j.jsat.2008.10.004

McKay, J. R., & Hiller-Sturmhofel, S. (2011). Treating alcoholism as a chronic disease: approaches to long-term continuing care. Alcohol Res Health, 33(4), 356-370.

National Committee on Quality Assurance. (2018). NCQA Updates Quality Measures for HEDIS 2018. from http://www.ncqa.org/newsroom/details/ncqa-updates-quality-measures-for-hedisreg-2018?ArtMID=11280&ArticleID=85&tabid=2659

Popovici, I., French, M. T., & McKay, J. R. (2008). Economic evaluation of continuing care interventions in the treatment of substance abuse: recommendations for future research. Eval Rev, 32(6), 547-568. doi: 10.1177/0193841X08316311

Reif, S., Acevedo, A., Garnick, D. W., & Fullerton, C. A. (2017). Reducing Behavioral Health Inpatient Readmissions for People With Substance Use Disorders: Do Follow-Up Services Matter? Psychiatr Serv, 68(8), 810-818. doi: 10.1176/appi.ps.201600339

Rubinsky, A. D., Ellerbe, L. S., Gupta, S., Phelps, T. E., Bowe, T., Burden, J. L., & Harris, A. H. S. (2017). Outpatient Continuing Care after Residential Substance Use Disorder Treatment in the U.S. Veterans Health Administration: Facilitators and Challenges. Subst Abus, 0. doi: 10.1080/08897077.2017.1391923

Schaefer, J. A., Ingudomnukul, E., Harris, A. H., & Cronkite, R. C. (2005). Continuity of care practices and substance use disorder patients´ engagement in continuing care. Med Care, 43(12), 1234-1241.

Tai, B., & Volkow, N. D. (2013). Treatment for substance use disorder: opportunities and challenges under the affordable care act. Soc Work Public Health, 28(3-4), 165-174. doi: 10.1080/19371918.2013.758975