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: Ctrl + click link 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 #: 0291

Corresponding Measures:

De.2. Measure Title: EMERGENCY TRANSFER COMMUNICATION MEASURE

Co.1.1. Measure Steward: University of Minnesota Rural Health Research Center

De.3. Brief Description of Measure: Percentage of all patients transferred from an Emergency Department to another healthcare facility whose medical record documentation indicated that all required information was communicated (sent) to the receiving facility within 60 minutes of transfer For all data elements, the definition of 'sent' includes the following:

- Hard copy sent directly with the patient, or
- Sent via fax or phone within 60 minutes of patient departure, or
- Immediately available via shared Electronic health record (EHR) or Health Information Exchange (HIE) (see definition below)

For purposes of this measure, a shared electronic health record (EHR) is defined as one where data entered into the system is immediately available at the receiving site. Facilities using the same EHR vendor or a Health Information Exchange (HIE) cannot assume immediate access by the receiving facility to the transferred patient's record.

1b.1. Developer Rationale: Improved timely communication of data elements will facilitate a better understanding of the patients´ condition prior to arrival at the receiving facility and reduce duplication of tests and procedures.

Improved timely communication of data elements will facilitate a better understanding of the patients´ condition prior to arrival at the receiving facility and reduce duplication of tests and procedures.

S.4. Numerator Statement: Numerator Statement: Number of patients transferred from an ED to another healthcare facility whose medical record documentation indicated that all of the following relevant elements were documented and communicated to the receiving hospital in a timely manner:

- Home Medications
- Allergies and Reactions
- Medications Administered in ED
- ED Provider Note

- Mental Status and Orientation Assessment
- Reason for Transfer and Plan of Care
- Tests and/or Procedures Performed
- Tests and/or Procedures Results

S.6. Denominator Statement: Denominator Statement: Transfers from an ED to another healthcare facility

Included Population: All transfers from an ED to another healthcare facility

Excluded Populations: Patients observation status.

S.8. Denominator Exclusions: All emergency department patients not discharged to another healthcare facility. Those admitted, sent home, left AMA, those on observations status, etc.

De.1. Measure Type: Process

S.17. Data Source: Claims, Electronic Health Data, Electronic Health Records, Management Data, Paper Medical Records

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Nov 15, 2007 **Most Recent Endorsement Date:** Sep 18, 2014

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? These Eight Data elements should be communicated for continuity of care and cost containment.

- Home Medications
- Allergies and Reactions
- Medications Administered in ED
- ED Provider Note
- Mental Status and Orientation Assessment
- Reason for Transfer and Plan of Care
- Tests and/or Procedures Performed
- Tests and/or Procedures Results

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria (“maintenance”). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

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 or **Summary of prior review in [year]**

- The Emergency Transfer Communication Measure captures the percentage of all patients transferred from an ED to another healthcare facility whose medical record documentation indicated that all required information was communicated (sent) to the receiving facility within 60 minutes of transfer. For all data elements, the definition of 'sent' includes the following:
  - Hard copy sent directly with the patient, or
  - Sent via fax or phone within 60 minutes of patient departure, or
  - Immediately available via shared electronic health record (EHR) or Health Information Exchange (HIE)
- All of the following must be captured in order to be included in the numerator:
  - Home Medications
  - Allergies and Reactions
  - Medications Administered in ED
  - ED Provider Note
  - Mental Status and Orientation Assessment
  - Reason for Transfer and Plan of Care
  - Tests and/or Procedures Performed
  - Tests and/or Procedures Results
- Please note: during the last review, this measure was granted an exception to evidence due to the lack of clear retrospective studies and the unethical nature of RCTs related to this topic.
- Developer provided an updated logic model depicting the connection between patient ED visit, care provided and documented, transfer occurring, and positive outcomes associated with having a record of the patient's ED visit travel with them to the next care setting
- Developer provides the rationale for reducing the number of data elements in the measure from 27 to 8.
  - In 2018, as part of the Rural Quality Improvement Technical Assistance (RQITA) program, Stratis Health, in partnership with the University of Minnesota Rural Health Research Center, convened a Technical Expert Panel (TEP) to review, revise, and update the EDTC measures and the related specifications manual.
  - The TEP recommended significant changes to help streamline and modernize the measure including reducing the total number of data elements from 27 to 8, updating the definition of 'sent' to better address communication via EHR or HIE, and clarifying specific definitions of individual data elements.

- Developer provided additional evidence on the vulnerability of patients being transferred and the opportunity for improvement on transfer communication.

Changes to evidence from last review

☒ **The developer provided updated evidence for this measure:**

Updates:

- Developer cites as evidence that "communication problems are a major contributing factor to adverse events in hospitals, accounting for 65% of sentinel events tracked by The Joint Commission."
- Developer does not provide sufficient evidence that the process dictated by the measure leads to better outcomes due to the absence of studies.

Questions for the Committee:

- What is the relationship of this measure to patient outcomes?
- How strong is the evidence for this relationship?
- Is the evidence directly applicable to the process of care being measured?
- Does the Committee believe that an exception to evidence is warranted?

Guidance from the Evidence Algorithm

From NQF "Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement" September 2019, pg 15:

Box 1 – Measure is not an outcome measure or PRO → Box 3 – Evidence is not based on a systematic review → Box 7 – No empirical evidence submitted → Box 10 – Possible outcome measures available (readmission to ER due to poor transfer of information) → No exception → Insufficient

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

RATIONALE: Measure is rated insufficient due to lack of existing evidence.

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.

- Results provided suggest a mean EDTC performance rate of 0.75 and 0.79 repectively in quarters 1 and 4 of 2017
- The interquartile range was approximately 0.65 – 0.97 for both quarters, indicating a substantial spread and continued opportunity for improvement.

Disparities

- Developer does not offer any disparities data.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- If no disparities information is provided, are you aware of evidence that disparities exist in this area of healthcare?

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

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

1a. Evidence: *For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission?For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.*

- Although the developer has provided a logic model, there still does not appear to be empirical support linking the transfer of patient information between facilities to patient outcomes
- Acceptable
- The developer provides updated information about evidence. My question is What evidence exists that records SENT are RECEIVED and REVIEWED? It seems to me that sending the information is important, but also is reviewing the information by the recipient.
- Low quality of evidence
- Updated evidence cited by the developer relates to relationship between the measure and adverse events in hospitals; however, the evidence doesn't include studies that demonstrate a linkage between the measure process and better outcomes.
- There is no stated evidence that links the process measure to improved outcomes.
- If the receiving facility does not have the ability to repeat the test then there is a tie to patient outcomes. Tests provide the data to inform doctors. This measure is disease agnostic and therefore inherently the relationship to patient outcomes will vary. Another issue is duplicate costs. It is assumed that the facility receiving the patient will re-do the tests. This is wasteful. Therefore, the evidence exists to support the measure focus. An exception to evidence is warranted because even if the evidence is lacking it does not mean the measure is not valid.
- No concerns
- Strength of evidence is low and has not been built on over the last three years which is quite disapppointing
- This is a process metric; no evidence based link to the outcome

1b. Performance Gap: *Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?*

- Based on the summary within patients (0,1) of the 8 new data elements at the facility level, there appears to be a demonstrable gap between deciles in the MBQIP data provided.
- No concerns
- This was rated as Moderate.
- Performance Gap demonstrated e.g. through the MN data cited. http://www.mnhealthscores.org/search/site//bundle/hospital/topics/931/#/results?topics=M931&viewmode=detail&page=13&non_rpt_hidden=y&columnname=M931&columntosort=M931&sortorder=desc No data on subgroups to demonstrate disparities, which raise the question of whether there should be adjustments for specific infrastructural and patient level differences
- There is opportunity for improvement in light of the EDTC performance rates (,75 and .79 for quarters 1 and 4 of 2017.
- Current data was provided. There was variability in the measured performance. No data related to disparities.
- Disparities data not provided; performance gap seems reasonably broad to warrant measurement

- Current performance data on the measure shows a performance rate of .75 to .79 with a big range of .65 to .97. A national performance measure could help create a case for low performing organizations to learn from the nearly perfect performing organizations. Clearly, there are processes that work for timely communication.
- Yes, there is a gap. There are no disparities data provided, as the developer states they do not exist. Though this may be OK for maintenance for now, we need to advocate for studies of disparities in this area moving forward.
- Gap identified
- Looking at transfer data, gap does exist; opportunity for improvement quanitfied; no inclusion to address national performance.
- Evidence

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

Scientific Acceptability: Preliminary Analysis Form

Measure Number: NQF 0291

Measure Title: Emergency Transfer Communication Measure

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**

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_0291" 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.**

   None identified.

RELIABILITY: TESTING

Submission document: "MIF_0291" 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**

   Submission document: Testing attachment, section 2a2.2

   - Data element reliability was assessed using Cohen's kappa score testing and interrater reliability testing.
   - Score level reliability was assessed using the beta-binomial methodology described by Adams (2009).

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   - Data element reliability was mixed, with fair to substantial crude agreement ranging from 69.2 – 81.7%, but low to moderate Kappa values ranging from 0.08 – 0.59 and a mean of 0.22.
   - Average signal to noise testing for two separate quarters was 0.95, indicating a high level of reliability at the score level.

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.**

    The agreement between abstractors at the data element level from the Cohen's kappa analysis weakens the confidence in the reliability of the measure at the data element level. The measure does demonstrate better crude agreement, and performs well at the score level.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: Testing attachment, section 2b2.

    No exclusions

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

    Submission document: Testing attachment, section 2b4.

    - Developer analyzes improvement of hospitals by type over time
    - The data include transfers to all facilities that provide clinically trained staff.
    - CAH to tertiary hospital communication has improved over time.
    - CAH transfers to non-acute facilities such as nursing homes, assisted living, detox centers still needs improvement.

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.
    No concerns.

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

    Submission document: Testing attachment, section 2b6.

    No concerns.

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**

The measure is a simple process measure entirely dictated by facility staff that should not be reasonably impacted by patient social risk factors**.**

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

    - Developer compared NQF 0291 with three ED quality performance measures that are related to process of care.
        - Influenza immunization rates of patients
        - Fibrinolytic therapy received in the ER within 30 minutes
        - Aspirin upon ER arrival
    - Developer reports Pearson correlation coefficients between the measures.
20. **Assess the results(s) for establishing validity**

    Submission document: Testing attachment, section 2b2.3

    - Results wear weakly positive and statistically significant, which was expected given the difference between communication processes and appropriate care processes.
        - Influenza measure range was 0.22 – 0.25 among the three EDTC domains
        - Fibrinolytic measure range was 0.06 – 0.08 among the three EDTC domains
        - Aspirin measure range was 0.16 – 0.17 among the three EDTC domains
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.**

Testing method was appropriate and results were only weakly positive, they were both significant and in the anticipated direction.

ADDITIONAL RECOMMENDATIONS

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.**

No additional recommendations.

Reliability

- Data element reliability was assessed using Cohen's kappa score testing and interrater reliability testing. Score level reliability was assessed using the beta-binomial methodology described by Adams (2009).
- Data element reliability was mixed, with substantial crude agreement ranging from 69.2 – 81.7% agreement, but low to moderate Kappa values ranging from 0.08 – 0.59 and a mean of 0.22.
- Average signal to noise testing for two separate quarters was 0.95, indicating a high level of reliability at the score level.
- The agreement between abstractors at the data element level from the Cohen's kappa analysis does weaken the confidence in the reliability of the measure at the data element level, but the measure does demonstrate reasonable crude agreement, and performs well at the score level.

Validity

- Developer compared NQF 0291 with three ED quality performance measures that are related to process of care.
    - Influenza immunization rates of patients
    - Fibrinolytic therapy received in the ER within 30 minutes
    - Aspirin upon ER arrival
- Developer reports Pearson correlation coefficients between the measures.
- Results were weakly positive and statistically significant, which was expected given the difference between communication processes and appropriate care processes.
    - Influenza measure range was 0.22 – 0.25 among the three EDTC domains
    - Fibrinolytic measure range was 0.06 – 0.08 among the three EDTC domains
    - Aspirin measure range was 0.16 – 0.17 among the three EDTC domains
- Testing method was appropriate and results were only weakly positive, they were both significant and in the anticipated direction.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The staff is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The staff is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

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

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

RATIONALE: Data element reliability scores for agreement among abstractors produced low-value kappas. While the score level testing was good, the method used assumes random rather than systematic error (abstractor agreement would be systematic).

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

2a1. Specifications: *Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?*

- The specifications appear adequate.
- No concerns
- There seems to be some problems with reliability at the data element level. This may need some exploration by the developer. I would like to see a response to this.
- Low inter-rater reliability suggests weak reliability. Measures do not include specific expectations regarding transmision of vitial signs, which is an important set of information promoted by ACEP

Practice policy of 2019
file:///C:/Users/dcase/OneDrive/IPO4Health/Casey%202020/NQF%202020/PEF%20Files/Feb%202020%20Call/0291%20Emergency%20Transfer%20Communication/transfer-of-patient-care-between-ems-providers-and-receiving-facilities%201-19.pdf

- Definitions seem clear, no concerns
- Low interrate reliability Kappa scores raise concerns for a complex process emasure requiring chart abstraction
- None identified
- Inter-rater reliability for individual elements was low. The developers provided a rationale for this (lack of support/opportunity for clarification for reviewers, due to lack of resources) but we should discuss this.
- Data elements reasonable
- NQF staff evaluated reliability as no; agree; these are yes no questions with a sigificant amount of subjectiveness

2a2. Reliability testing: *Do you have any concerns about the reliability of the measure?*

- The impact of the exclusion of facilities with <45 patients per quarter does not appear to be addressed. Inter-rater reliability appears to remain poor.
- None
- I would be interested in seeing a higher rating for reliability based on the the measure taken in total v. the data elements.
- Yes as above
- Mixed results, yields low confidence level. The agreement between abstractors at the data element level from the Cohen's kappa analysis weakens the confidence in the reliability of the measure at the data element level. However, the measure demonstrates better crude agreement, and performs well at the score level.
- If I'm interpreting correctly- there's inconsistency in how abstractors are interpreting and reporting out data for their hospital. If that is true- I am concerned about using this as performance measure.
- Score level assessments reassuring but low Kapp scores concerning
- Data element reliability was mixed, with fair to substantial crude agreement ranging from 69.2 – 81.7%, but low to moderate Kappa values ranging from 0.08 – 0.59 and a mean of 0.22. Would like some perspective from reviewing team on how this compares to other measures reviewed. I understand the staff is satisfied with the reliability testing for the measure and do not want to require a vote based on my first time on the team.
- See last comment re IRR for elements
- Feasible
- Yes - staff rated as low but mentioned performs well at the score level

2b1. Validity testing: *Do you have any concerns with the testing results?*

- The associations between the measure and validation variables (i.e. immunization screening), although some are statistically significant, show less than 5% shared variance.
- No concerns.
- Some as it relates to low reliability.
- Unable to comment. The rationale for conducting a comparison between communication processes and appropriate care processes is unclear to me.

- would have preferred to see comparison to outcome measure results/clinical outcomes, rather than other process measures, especially given the low correlations (although directionally supportive of association)
- Support NQF findings of no concerns and will support experienced reviewers need for a discussion.
- Validity weak
- Staff rated as moderate. I do have concerns with the yes/no format, e.g. med list - if not available not rated; does not help with the outcome of having an accurate med list

Validity- Threats to Validity: *Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) 2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?*

- There does not appear to be any analysis of the impact of excluding hospitals with <45 patients per quarter. Also, it is unclear whether hospitals could participate in some quarters but not others should their patient transfers exceed 45.
- No concerns
- Same as above related to lack of adjustment for infrastrctural and patient characteristics
- Developer analyzes improvement of hospitals by type over time. Data include transfers to all facilities that provide clinically trained staff; Critical Access Hospital (CAH) to tertiary hospital communication has improved over time; and CAH transfers to non-acute facilities such as nursing homes, assisted living, detox centers still needs improvement. No other threats to validity noted.
- Yes. If the data elements are sent, but for some reason, not received by the receiving facility, the post-discharge communication is still broken.
- They do show a relationship between sample volume/response rate and reliability - low response rates would be a threat to validity
- Support NQF findings of testing method was appropriate and results were only weakly positive, they were both significant and in the anticipated direction.
- Results were weakly positive and statistically significant, which was expected given the difference
- Moderate rating; no concerns

Other Threats to Validity: *Other Threats to Validity (Exclusions, Risk Adjustment)2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure?2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?*

- No risk adjustment is proposed/ indicated/ required.
- Yes
- I wondered about exclusion of dismissal to home with Home Health. Shouldn't Home Health also receive this information in either a written or electronic form on a timely basis?
- As above previuosly mentioned challenges with lack of adjustment methods
- No; N/A
- measure not risk-adjusted
- No concerns
- No RA; exclusions reasonable

2c. Composite Performance Measure: *Composite Analysis (if applicable): Do analyses demonstrate the component measures fit the quality construct and add value? Do analyses demonstrate the aggregation and weighting rules fit the quality construct and rationale?*

- No data are provided on the contribution of the 8 individual components to the overall score for this measure. Using a formative measurement model, these data could have been provided in the reliability section.
- Yes
- Yes, prior submissions were of individual components and measure developers were instructed to create composite. But now lacking specific information on vital signs deemed important in ACEP Practice Policy file:///C:/Users/dcase/OneDrive/IPO4Health/Casey%202020/NQF%202020/PEF%20Files/Feb%202020%20Call/0291%20Emergency%20Transfer%20Communication/transfer-of-patient-care-between-ems-providers-and-receiving-facilities%201-19.pdf
- 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 health records.
- Hospitals with electronic medical records shared with the patient´s destination can capture these elements easily. Hospitals who do not share electronic medical records with the destination hospital require additional paper trails for capture of necessary process elements.
- No licensure fees required to use the measure.

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?
- Is the data collection strategy ready to be put into operational use?

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

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility: *Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?*

- Unclear
- No concerns
- Especially with the trimming of criteria from 27 to 8 elements, this is a Feasible measure.
- Some facilities may be at a disadvantage if incomplete EHR implementation and interoperability constraints with other facilities.
- No concerns about data collection and undue burden, thus, feasibility seems relatively high.

- High burden process measure requiring manual abstraction/documentation
- This measure is feasible. The required data elements are routinely generated and used. When patients transfer their data needs to go with them.
- Highly feasible
- Would rate as moderate; although the questions are easy to answer; there is effort to get the most accruate information for the receiving facility; med list again would be an example; the transferring faciity would need to put forth effort to get the list.

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

Public Reporting

- Program Name: Minnesota Statewide Quality Reporting and Measurement System.
  - http://www.mnhealthscores.org/hospital-quality-patient-experience-ratings

Quality Improvement (external benchmarking to organizations)

- MN statewide forms
  - http://www.health.state.mn.us/healthreform/measurement/index.html
- Medicare Beneficiary Quality Improvement Project (MBQIP)
  - https://www.ruralcenter.org/tasc/mbqip

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

- Reports are produced by a contractor to HRSA FORHP on a quarterly basis. Each quarter, the FORHP Project Officer for each state Flex program distributes the individual hospital and state summary data

to the individual state Flex programs, who then in turn distribute the reports to the CAHs in their state. Individual state Flex programs provide support to CAHs on interpretation and use of the measure, but a guide on how to use the reports is also available. MN CAHs receive their data through this same process even though the measure is also used for public reporting through the MN SQRMS program.

- Voluntary participants appreciate the feedback and have made simple process design improvements. Non-voluntary participants struggle with the indications of inadequate process completion.
- Non-users have not been available to the measure sponsors. Some ideas may be that because CAHs have limited resources, their quality measurement resources may be focused on other issues. Some CAH may have improved their scores on these measures to near 100% and shifted their focus on other issues.

Additional Feedback: N/A

***Questions for the Committee**:*

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

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

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

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- Improvement: Data and Trends (Q1 2015 – Q3 2016)
- Every EDTC sub-measure and the composite EDTC-All measure has consistently improved between Q1
- 2015 and Q3 2016. Most markedly, EDTC-All has risen from 51.8% to 74.3% (a 22.5 percentage point increase).

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 [unexpected findings]

Potential harms: None identified

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**

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

4a. Use: *4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is*

measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided? 4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- Unclear.
- Yes
- This has high use and usability
- Currently used publicly in MN. Unclear if measure is impacting clinical outcomes.
- "Developer indicates the measure is being reported as follows:
    - Minnesota Statewide Quality Reporting and Measurement System. http://www.mnhealthscores.org/hospital-quality-patient-experience-ratings
    - Quality Improvement (external benchmarking to organizations) MN statewide forms http://www.health.state.mn.us/healthreform/measurement/index.html
    - Medicare Beneficiary Quality Improvement Project (MBQIP) https://www.ruralcenter.org/tasc/mbqip
    - Quality Improvement (Internal to the specific organization) Medicare Beneficiary Quality Improvement Project (MBQIP) https://www.ruralcenter.org/tasc/mbqip. Use feedback provided by the developers include:
    - Reports are produced by a contractor to HRSA FORHP on a quarterly basis. Each quarter, the FORHP Project Officer for each state Flex program distributes the individual hospital and state summary data to the individual state Flex programs, who then in turn distribute the reports to the CAHs in their state.
    - Individual state Flex programs provide support to CAHs on interpretation and use of the measure, but a guide on how to use the reports is also available. MN CAHs receive their data through this same process even though the measure is also used for public reporting through the MN SQRMS program.
    - Voluntary participants appreciate the feedback and have made simple process design improvements. Non-voluntary participants struggle with the indications of inadequate process completion.
    - Non-users have not been available to the measure sponsors. Some ideas may be that because CAHs have limited resources, their quality measurement resources may be focused on other issues. Some CAHs may have improved their scores on these measures to near 100% and shifted their focus on other issues.
    - Reports are produced by a contractor to HRSA FORHP on a quarterly basis. Each quarter, the FORHP Project Officer for each state Flex program distributes the individual hospital and state summary data to the individual state Flex programs, who then in turn distribute the reports to the CAHs in their state.
    - Individual state Flex programs provide support to CAHs on interpretation and use of the measure, but a guide on how to use the reports is also available. MN CAHs receive their data through this same process even though the measure is also used for public reporting through the MN SQRMS program.
    - Voluntary participants appreciate the feedback and have made simple process design improvements. Non-voluntary participants struggle with the indications of inadequate process completion.
    - Non-users have not been available to the measure sponsors. Some ideas may be that because CAHs have limited resources, their quality measurement resources may be focused on other issues. Some CAHs may have improved their scores on these measures to near 100% and shifted their focus on other issues."

- Yes- those being measured have been offered an opportunity to provide feedback. It's not clear to me whether their input has been incorporated into measure changes.
- In current use and shows significant performance improvements over time
- Based on the feedback from voluntary participants according to the NQF report they appreciate the feedback and have made simple process design improvements. This shows that the non-voluntary participants and the non-users may have an opportunity to further the goal of high-quality, efficient healthcare through learning. With the results to date, the measure been vetted in real-world settings by those being measured and demonstrates performance results for both accountability and performance improvement activities.
- No concerns
- Usable
- Need more info on the non-users; difficult to assess without that infomration; users have made some imporvements based on the data

4b. Usability: *4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.*

- High Usability
- Unclear without additional information
- "Based on trend data presented, ongoing tracking and monitoring can be used to assess adherence to protocol.
    - Data and Trends (Q1 2015 – Q3 2016)
    - Every EDTC sub-measure and the composite EDTC-All measure has consistently improved between Q1
    - 2015 and Q3 2016. Of significance- EDTC-All has risen from 51.8% to 74.3% (a 22.5 percentage point increase). Based on information provided, no potential harms were noted. It appears the benefits of this 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)."
- No unintended consequences.
- I have concerns both about burden to collect/document measure results and the absence of disparities data - I do not think the developer's rationale that there is no rationale for disparities signifies a substantive argument for not providing these data; especially for a measure that is asking rural CAHs to sbmit documentation, I think it would be very important to understand if performance differs by patient-level or community-level factors of social risk
- The measure is usable and needed. However, as a process measure it could be burdensome to productivity. If transfer of information was automated, there would be no need to measure time of transfer. Is there a way to make this a pass/fail. In other words, information was here when needed or not. Does the information always have to be there in 60 minutes? Maybe the patient is still being transferred. I think this is an opportunity to look at process flow and determine if there is a consistent way to streamline instead of added more measures.
- No concerns
- Usable
- Would rate as moderate; could be more meangful with feedback from non-users and direct link to outcome

Criterion 5: Related and Competing Measures

Related or competing measures
No measures identified.
Harmonization
N/A

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

Related and Competing: *Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?*

- No competing measures, BUT, it would be good to see a related measure that the information that is received by the Facility accepting the Transfer is actually REVIEWED.
- Support NQF findings.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 2/7/2020

- No NQF members have submitted a support/non-support choice
- No NQF members have commented

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

NQF_evidence_attachment_Sep2017Evidence_-subcriterion___10.22.19.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.

Yes

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0291

Measure Title: Emergency Department Transfer Communication Measure

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 11/23/2019

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

Outcome

☐ Outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the 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*): Click here to name the intermediate outcome

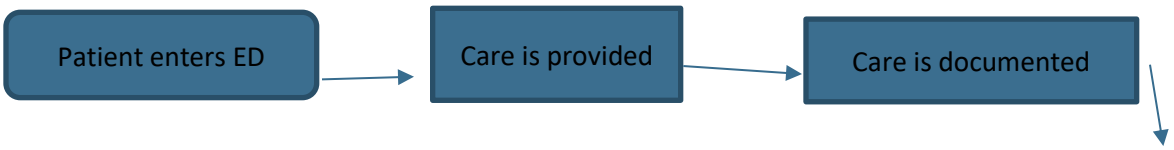
☒ Process: documentation of sending information to a receiving facility

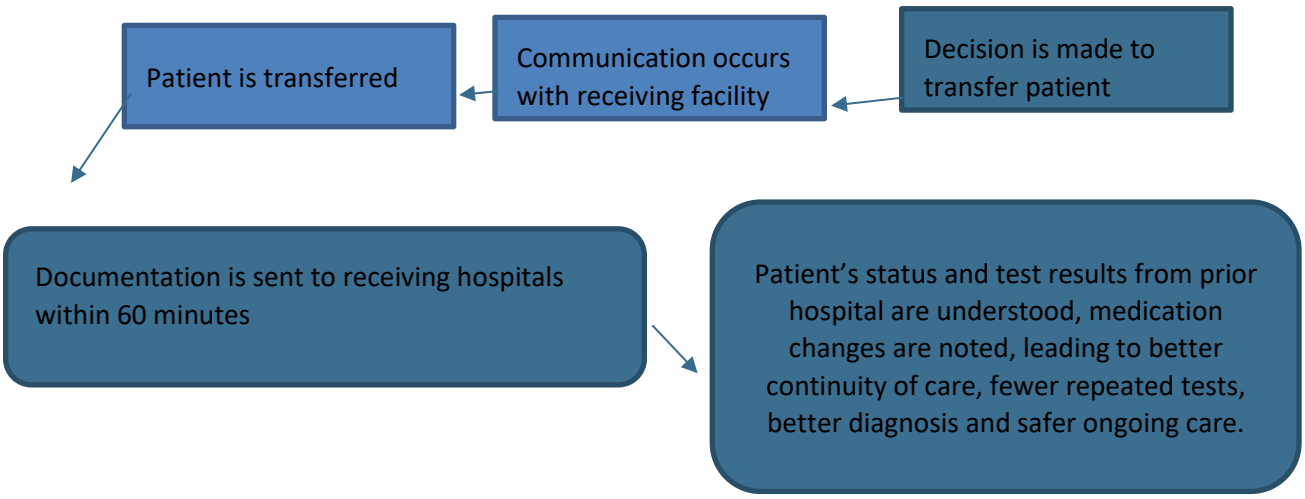
☒ Appropriate use measure: documentation of info sent.

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

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





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

****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.

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.

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.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure.

2019 additions to background literature support

There is widespread agreement that patients receiving care from multiple providers across the care spectrum are at risk for poor quality care, adverse events, higher expenditures and incomplete medication management. There are no randomized control studies that directly examine the impact of structured communication between Critical Access Hospitals (CAHs) that transfer patients to facilities for other acute and non-acute care. Studies demonstrate that transfer patients experience higher cost of care and poorer outcomes

"Patient transfers between hospitals are becoming more common in the United States. Disease-specific studies have reported varying outcomes associated with transfer status. However, even as national quality improvement efforts and regulations are being actively adopted, forcing hospitals to become financially accountable for the quality of care provided, surprisingly little is known about transfer patients or their outcomes at a population level." (Hernandez-Boussard et al., 2017)

Communication, efficiency and appropriateness are key factors advanced as impacting the quality and safety of non-emergency transport services. The safety of transferred patients is sometimes compromised by poor standardization and failures in the communication process. (Hains et al., 2011) Patient care during inter-facility transfer depends not only on the expertise provided by the receiving facility, but also on timely and accurate patient information received from the transferring institution.(Szary et al., 2010)

In a large national study of academic health center 2011-2012 discharges inter hospital transfer (IHT) patients are at a higher risk of inpatient mortality after controlling for patient characteristics and risk of mortality measures (odds ratio: 1.36, 95% confidence interval: 1.29–1.43).(Sokol-Hessner et al., 2016)

Transfer patients in the lowest risk group have longer length of hospital stay and ICU stays. (Golestanian et al., 2007) They cost more solely due to their transfer status. What about being transferred sets them up for these higher costs? Could it be communication with the transferring hospital?

Transfer patients use more resources and have worse outcomes than nontransfer patients. They have statistically significant longer length of stays (13 to 4.5), more non-routine dispositions (53% to 68%), higher risk-adjusted inpatient mortality (4.6% to 2.1%), and higher risk-adjusted Patient Safety Indicators (PSI) In 4 of 5 indicators. Most CMS sponsored hospital quality measures do not include transferred patients. "Carefully constructed national quality measures for transfer of care should be designed and validated."(Hernandez-Boussard et al., 2017)

"This expanding evidence base demonstrates that serious deficiencies in quality exist for patients undergoing transitions across sites of care. Qualitative studies produced consistent results, demonstrating that patients are often unprepared for their self-management role in the next care setting, receive conflicting advice regarding chronic illness management, are often unable to reach an appropriate health care practitioner who has access to their care plan when questions arise, and have minimal input into their care plan. Quantitative studies documented that quality and patient safety are compromised during the vulnerable period when patients transition between different settings

because of high rates of medication errors, incomplete or inaccurate information transfer, and lack of appropriate follow-up care. During care transitions, patients receive medications from different prescribers, who rarely have access to patients' comprehensive medication lists. Collectively, these types of problems conspire to increase rates of recidivism to high intensity care settings when patients' care needs are not met, leading to greater health care costs." (Coleman et al., 2006)

Hernandez-Boussard et al., 's population-wide study provides timely analyses of the characteristics of this particularly vulnerable and sizable inpatient population. Opportunity exists for improvement on these processes." (Hernandez-Boussard et al., 2017)

Szary's prospective study quantified compliance with inter-facility transfer communication and revealed an opportunity for improvement. Introduction of a simple written template to enhance communication between providers improved the quality of transfer information. (Szary et al., 2010)

From Previous Submission

Background of the Measure: In 2003, an expert panel convened by the University of Minnesota Rural Health Research Center and Stratis Health identified ED care as an important quality assessment measurement category for rural hospitals. While emergency care is important in all hospitals, it is particularly critical in rural hospitals where the size of the hospital and geographic realities make organizing triage, stabilization, and transfer of patients more important.

In 2018, as part of the Rural Quality Improvement Technical Assistance (RQITA) program, Stratis Health, in partnership with the University of Minnesota Rural Health Research Center, convened a Technical Expert Panel (TEP) to review, revise, and update the EDTC measures and the related specifications manual. The Panel members represented national experts in hospital ED physicians and nurses, quality measurement, electronic health records, and data analytics. The TEP met three times via conference call to review the measure specifications and discussion was framed around three primary issues and challenges including EDTC in a "wired" world, appropriate population for transfers, and clinical relevance of specific data elements. The TEP recommended significant changes to help streamline and modernize the measure including reducing the total number of data elements from 27 to 8, updating the definition of 'sent' to better address communication via EHR or HIE, and clarifying specific definitions of individual data elements.

Communication between providers promotes continuity of care and may lead to improved patient outcomes. These measures were piloted by rural hospitals in Hawaii, Iowa, Maine, Minnesota, Missouri, Nebraska, Nevada, New York, Ohio, Oklahoma, Pennsylvania, Utah, Washington, West Virginia, Wisconsin, and Wyoming; projects took place from October 2005 through July 2014. Results of the pilot projects indicated room for improvement in ED care and transfer communication. Aggregate project results are available at http://www.flexmonitoring.org/wp-content/uploads/2014/02/ds8.pdf and http://flexmonitoring.org/documents/FlexDataSummaryReport3.pdf.

This measure is being implemented by critical access hospitals (CAHs) in the Medicare Beneficiary Quality Improvement Project (MBQIP) because small rural hospitals frequently transfer a higher proportion of emergency department (ED) patients than larger urban facilities. It is an important goal of MBQIP to help hospitals improve care transitions, including ED transfers, in order to reduce preventable hospital readmissions and adverse events in hospitals. Currently, 86% of the 1,318 CAHs participating in MBQIP are reporting EDTC-All.

In 2014, over 100 CAHs across eight states participated in a one-year special innovation project through the Centers for Medicare and Medicaid Services (CMS) led by Stratis Health. A case study discussing implementation of EDTC in Minnesota was also done by the Flex Monitoring Team (FMT), a Federal Office of Rural Health Policy (FORHP) funded consortium of research centers, in 2014. EDTC became a required MBQIP measure in 2015, and reporting rates among CAHs nationwide have risen dramatically since that time. Every EDTC sub-measure and the composite EDTC-All measure has consistently improved between Q1 2015 and Q3 2016. Most markedly, EDTC-All has risen from 51.8% to 74.3% (a 22.5 percentage point increase). Reporting rates and measure scores have improved.

Communication problems are a major contributing factor to adverse events in hospitals, accounting for 65% of sentinel events tracked by The Joint Commission. In addition, research indicates that deficits exist in the transfer of patient information between hospitals and primary care physicians in the community, and between hospitals and long-term facilities. Transferred patients are excluded from the calculation of most national quality measures, such as those used in Hospital Compare. The Hospital Compare Web site was created to display rates of Process of Care measures using data that are voluntarily submitted by hospitals.

The Joint Commission has adopted National Patient Safety Goal 2, "Improve the Effectiveness of Communication Among Caregivers." This goal required all accredited hospitals to implement a standardized approach to hand- off communications, including nursing and physician handoffs from the emergency department (ED) to inpatient units, other hospitals, and other types of health care facilities. The process must include a method of communicating up-to-date information regarding the patient's care, treatment, and services; condition; and any recent or anticipated changes. (Note: The National Patient Safety Goals are reviewed and modified periodically. In 2013 a communication goal focuses on the communication of test results.) http://www.jointcommission.org/assets/1/6/HAP_NPSG_Chapter_2014.pdf Limited attention has been paid to the development and implementation of quality measures specifically focused on patient transfers between EDs and other health care facilities. Examples are patients transferred between an ED and a skilled nursing facility with their often vulnerable and fragile populations. These measures are important for all health care facilities, but especially so for small rural hospitals that transfer a higher proportion of ED patients.

While many aspects of hospital quality are similar for urban and rural hospitals (e.g., providing heart attack patients with aspirin), the urban/rural contextual differences result in differences in emphasis on quality measurement. Because of its role in linking residents to urban referral centers, important aspects of rural hospital quality include triage-and-transfer decision making about when to provide a particular type of care, transporting patients, and coordinating information flow to specialists beyond the community. Because of their size, rural hospitals are less likely to be able to provide more specialized services, such as cardiac catheterization or trauma surgery. Rural residents often need to travel greater distances than urban residents to get to a hospital initially. In addition, their initial point of contact is less likely to have specialized services and staff found in tertiary care centers, so they are also more likely to be transferred. These size and geographic realities increase the importance of organizing triage, stabilization, and transfer in rural hospitals which, in turn, suggest that measurement of these processes is an important issue for rural hospitals.

The ED Transfer Communication measure aims to provide a means of assessing how well key patient information is communicated from an ED to any healthcare facility. They are applicable to patients with a wide range of medical conditions (e.g., acute myocardial infarction, heart failure, pneumonia, respiratory compromise and trauma) and are relevant for both internal quality improvement purposes and external reporting to consumers and purchasers. The results of the field tests suggest that significant opportunity exists for improvement on these processes.

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

Routine PubMed Searches.

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

New references for 2019 submission

Coleman, E. A., Parry, C., Chalmers, S., & Min, S.-J. (2006). The care transitions intervention: Results of a randomized controlled trial. *Archives of Internal Medicine*, *166*(17), 1822–1828. https://doi.org/10.1001/archinte.166.17.1822

Golestanian, E., Scruggs, J. E., Gangnon, R. E., Mak, R. P., & Wood, K. E. (2007). Effect of interhospital transfer on resource utilization and outcomes at a tertiary care referral center. *Critical Care Medicine*, *35*(6), 1470–1476. https://doi.org/10.1097/01.CCM.0000265741.16192.D9

Hains, I. M., Marks, A., Georgiou, A., & Westbrook, J. I. (2011). Non-emergency patient transport: What are the quality and safety issues? A systematic review. *International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care*, *23*(1), 68–75. https://doi.org/10.1093/intqhc/mzq076

Hernandez-Boussard, T., Davies, S., McDonald, K., & Wang, N. E. (2017). Interhospital Facility Transfers in the United States: A Nationwide Outcomes Study. *Journal of Patient Safety*, *13*(4), 187–191. https://doi.org/10.1097/PTS.0000000000000148

Sokol-Hessner, L., White, A. A., Davis, K. F., Herzig, S. J., & Hohmann, S. F. (2016). Interhospital transfer patients discharged by academic hospitalists and general internists: Characteristics and outcomes. *Journal of Hospital Medicine*, *11*(4), 245–250. https://doi.org/10.1002/jhm.2515

Szary, N. M., Sarwal, A., Boshard, B. J., & Hall, L. W. (2010). Transfer of care communication: Improving communication during inter-facility patient transfer. *Missouri Medicine*, *107*(2), 127–130.

Brief article review of new citations: 2019

For each study you review, state the study design, report sample size and report the main result.

1. Szary NM, Sarwal A, Boshard BJ, Hall LW. Transfer of Care Communication: Improving Communication during Inter-Facility Patient Transfer. Missouri Medicine March/April 2010 102:2

Study Design: Pre Post prospective study quantified compliance with inter-facility transfer communication requested information. Intervention was a template of requested information used by the receiving hospital when contacted by the sending hospital. Observations did not differentiate between sources of transfers.

Sample Size: 60 pre observations, 43 post observations. Observations of information provided by sending facilities were made by resident physicians, at the University of Missouri 223 bed hospital in Columbia Missouri, who were receiving those transfers.

Main Result: Significant improvement was seen in the reception of Medication lists (P<0.01), and Imaging (P<0.01). Improvement in Lab data was documented (P<0.11). No significant increase was observed in discharge summaries.

Limitations: Small sample.

2. Golestanian E[1], Scruggs JE, Gangnon RE, Mak RP, Wood KE.

Effect of interhospital transfer on resource utilization and outcomes at a tertiary care referral center. Crit Care Med. 2007 Jun;35(6):1470-6.

Design: Observational cohort study. Comparing Acute Physiology and Chronic Health Evaluation (APACHE) III score, actual and predicted ICU and hospital lengths of stay, actual and predicted ICU and hospital mortality, and costs per admission.

Sample Size:4569 patients admitted to a tertiary ICU from 1997 to 2000.

Main Result: Crude comparison of directly admitted and transfer patients revealed that transfer patients had significantly higher APACHE III scores (mean, 60.5 vs. 49.7, p < .001), ICU mortality (14% vs. 8%, p < .001), and hospital mortality (22% vs. 14%, p < .001). Transfer patients also had longer ICU lengths of stay (mean, 6.0 vs. 3.8 days, p < .001) and hospital lengths of stay (mean, 20 vs. 15.9 days, p < .001). Stratified by disease severity using the APACHE III model, there was no difference in either ICU or hospital mortality between the two populations. However, in the transfer group with the lowest predicted mortality of 0-20%, ICU and hospital lengths of stay were significantly higher. In crude cost analysis, transfer patients' costs were $9,600 higher per ICU admission compared with nontransfer patients (95% confidence interval, $6,000-$13,400). Risk stratification revealed that the higher per-patient cost was entirely confined to the transfer patients with the lowest predicted mortality.

3. Interhospital Facility Transfers in the United States: A Nationwide Outcomes Study

Tina Hernandez-Boussard, PhD, MPH, MS*, Sheryl Davies, BA†, Kathryn McDonald, MA†, and N. Ewen Wang, M

J Patient Saf . 2017 December ; 13(4): 187–191. doi:10.1097/PTS.0000000000000148

Study Design: Retrospective observational study of in-hospital adverse events and discharge deposition for transfer patients vs nontransfer patients using Nationwide Inpatient Sample (NIS).

Sample Size: 1,397,712 transfer patients and 31,692,211 nontransfer patients for a total of 33,089,923 patients.

Main Result: We identified 1,397,712 transfer patients and 31,692,211 nontransfer patients. Age, sex, race, and payer were significantly associated with odds of transfer ($P < 0.05$). Transfer patients had higher risk-adjusted inpatient mortality (4.6 versus 2.1, $P < 0.01$), longer length of stay (13.3 versus 4.5, $P < 0.01$), and fewer routine disposition discharges (53.6 versus 68.7, $P < 0.01$). In-hospital adverse events were significantly higher in transfer patients compared with nontransfer patients ($P < 0.05$). With hospital accountability and value-based payments constituting an integral part of health care reform, documenting the quality of care delivered to transfer patients is essential before accurate quality assessment improvement efforts can begin in this patient population.

Our results suggest that transfer patients have inferior outcomes compared with nontransfer patients. Although they are clinically complex patients and assessing accountability as between the transferring and receiving hospitals is methodologically difficult, transfer patients must nonetheless be included in quality benchmark data to assess the potential impact this population has on hospital outcome profiles. With hospital accountability and value-based payments constituting an integral part of health care reform, documenting the quality of care delivered to transfer patients is essential before accurate quality assessment improvement efforts can begin in this patient population.

4. Interhospital Transfer Patients Discharged by Academic Hospitalists and General Internists: Characteristics and Outcomes

Lauge Sokol-Hessner, MD1*, Andrew A. White, MD2, Katherine F. Davis, RN3, Shoshana J. Herzig, MD, MPH1, Samuel F. Hohmann, PhD3

Journal of Hospital Medicine Vol 11 | No 4 | April 2016

Study Design: Retrospective cohort study that compares interhospital transfers (IHT) with those admitted from the emergency department to an academic health system.

Sample Size: 885,392 adult inpatients discharged from 158 academic health centers and affiliated hospitals from April 1, 2011 to March 31,2012.

Main Result: IHT patients are at a higher risk of inpatient mortality after controlling for patient characteristics and risk of mortality measures.

5. The Care Transitions Intervention

Results of a Randomized Controlled Trial

Eric A. Coleman, MD, MPH; Carla Parry, PhD, MSW; Sandra Chalmers, MPH; Sung-joon Min, PhD

ARCH INTERN MED/VOL 166, SEP 25, 2006 WWW.ARCHINTERNMED.CO

Study Design: Randomized control trial from September 1 2002 to Aug 31 2003 of patients moving across different health care settings who received communication tools, assertiveness encouragement, and "transition" coaching.

Sample Size: 750 community dwelling adults over age 65 who were admitted to the study hospital for 1 of 11 conditions.

Main Result: Coaching for chronically ill elders to ensure that their needs are met may reduce rates of subsequent rehospitalization.

6. Non-emergency patient transport: What are the quality and safety issues? A systematic review. Hains, I. M., Marks, A., Georgiou, A., & Westbrook, J. I. (2011). *International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care*, *23*(1), 68–75. https://doi.org/10.1093/intqhc/mzq076

   Study design: Systematic Review

   Sample: Twelve articles from 7 countries.

Main Results: Communication, efficiency and appropriateness are key factors that are advanced as impacting on the quality and safety of non-emergency transport services. Standardization of the non-emergency transport process shows promise in reducing risk and increasing efficiency. Applying information and communication technology to improve the quality of transport services has received little attention despite its potential benefits. Patient outcomes in relation to quality and safety of transport services are rarely measured. Available evidence suggests that safety of non-emergency patient transfers is sometimes compromised due to poor standards.

References from previous submission:

Baldwin LM, MacLehose RF, Hart LG, Beaver SK, Every N, Chan L. Quality of care for acute myocardial infarction in rural and urban US hospitals. J Rural Health 2004 Spring;20(2):99-108.

Cortes TA, Wexler S, Fitzpatrick JJ. The transition of elderly patients between hospitals and nursing homes. Improving nurse-tonurse communication. J Gerontol Nurs 2004 Jun;30(6):10-5; quiz 52-3. [5 references]

Ellerbeck EF, Bhimaraj A, Perpich D. Organization of care for acute myocardial infarction in rural and urban hospitals in Kansas. J Rural Health 2004 Fall;20(4):363-7.

Joint Commission on Accreditation of Healthcare Organizations. Sentinel events statistics. [Internet]. [Accessed 2007 Jul 18].

Klingner J, Moscovice I, Washington Rural Healthcare Quality Network and Stratis Health, Minnesota Quality Improvement Organization. Rural hospital emergency department quality measures: aggregate data report. Minneapolis (MN): University of Minnesota, Division of Health Services Research & Policy; 2007 Mar. 12 p. (Flex Monitoring Team data summary report; no. 3).

Klingner J, Moscovice I. Development and testing of emergency department patient transfer communication measures. J Rural Health 2012 Jan;28(1):44-53. [16 references]

Kripalani S, Lefevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. JAMA 2007 Feb 28;297(8):831-41. [133 references]

Newgard CD, McConnell KJ, Hedges JR. Variability of trauma transfer practices among non-tertiary care hospital emergency departments. Acad Emerg Med 2006 Jul;13(7):746-54.

University of Minnesota Rural Health Research Center, Stratis Health (Minnesota's Quality Improvement Organization), HealthInsight (Nevada and Utah's Quality Improvement Organization). Refining and field testing a relevant set of quality measures for rural hospitals. Final report submitted to the Centers for Medicare & Medicaid Services under contract no. 500-02-MN01. Bloomington (MN): Stratis Health; 2005 Jun 30. US Department of Health and Human Services. Hospital Compare Web site. [Web site]. [Accessed 2011 Feb 25].

Wakefield DS, Ward M, Miller T, Ohsfeldt R, Jaana M, Lei Y, Tracy R, Schneider J. Intensive care unit utilization and interhospital transfers as potential indicators of rural hospital quality. J Rural Health 2004 Fall;20(4):394-400.

Westfall JM, Van Vorst RF, McGloin J, Selker HP. Triage and diagnosis of chest pain in rural hospitals: implementation of the ACI-TIPI in the High Plains Research Network. Ann Fam Med 2006 Mar-Apr;4(2):153-8.

http://www.stratishealth.org/providers/ED_Transfer_Resources.html

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.

Improved timely communication of data elements will facilitate a better understanding of the patients´ condition prior to arrival at the receiving facility and reduce duplication of tests and procedures.

Improved timely communication of data elements will facilitate a better understanding of the patients´ condition prior to arrival at the receiving facility and reduce duplication of tests and procedures.

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.*

The validity and reliability testing reflect data on the new specifications. The new specifications are a subset of the old specifications, the analysis for validity and reliability contains only the 8 current elements.

Data from the MBQIP submission site from 2017 from over 1000 CAH hospitals is summarized here. It provides

the descriptive stats for 2 quarters in 2017 from the previous specifications. Data from the new specifications will be available after Oct 2020. The overall mean for all reported increased from 75% to 78 %, the standard error narrowed from .8% to .7%; both of these are statistically significant with over 1100 observations. Scores by decile show a swing to higher scores. In the approximate top decile, (top 9 %), there were 38% in Q1 and 42% in Q 4. Improvements in completion are being made over a short amount of time.

Descriptive stats
q1 2017 q4 2017

Mean 0.750450484 Mean 0.786365037
Standard Error 0.008025333 Standard Error 0.007349871
Median 0.845299145 Median 0.87755102
Mode 1 Mode 1
Standard Deviation 0.273568515 Standard Deviation 0.253010691
Sample Variance 0.074839732 Sample Variance 0.06401441
Range 1 Range 1
Minimum 0 Minimum 0
Maximum 1 Maximum 1
Sum 872.0234623 Sum 931.8425689
Count 1162 Count 1185

Decile approximations 2017 MBQIP
q1 Cumulative % q4 Cumulative %
0.00 2.84% 0.00 2.70%
0.41 13.43% 0.41 10.21%
0.65 26.85% 0.65 21.43%
0.76 39.67% 0.76 33.25%
0.82 47.33% 0.82 41.27%
0.88 55.77% 0.88 50.38%
0.91 62.82% 0.91 58.82%
0.94 68.33% 0.94 64.89%
0.97 74.53% 0.97 71.81%
More 100.00% More 100.00%

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.

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.*

No disparities data is available at this time.

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

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):*

Care Coordination, Care Coordination : Readmissions, Care Coordination : Transitions of Care, Safety, Safety : Medication

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 new specifications are only available to hospitals participating in the pilot in June 2019. Here is the link to the new specifications. http://www.stratishealth.org/documents/EDTC-Data-Specs-Manual-2019.pdf

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

This is not an eMeasure **Attachment:**

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

No data dictionary **Attachment:**

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

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

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

Not an instrument-based measure

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

Yes

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.

The Denominator has been clarified to include transfers to non-acute facilities such as nursing homes and to exclude patients in observation status. Continuity of care as derived from these data elements can prevent errors. The measure has been redesigned under the direction of a technical expert panel to include 8 of the original data elements.

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).

Numerator Statement: Number of patients transferred from an ED to another healthcare facility whose medical record documentation indicated that all of the following relevant elements were documented and communicated to the receiving hospital in a timely manner:

- Home Medications
- Allergies and Reactions
- Medications Administered in ED
- ED Provider Note
- Mental Status and Orientation Assessment
- Reason for Transfer and Plan of Care
- Tests and/or Procedures Performed
- Tests and/or Procedures Results

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).

Numerator Statement: Number of patients transferred from an ED to another healthcare facility whose medical record documentation indicated that all of the following relevant elements were documented and communicated to the receiving hospital in a timely manner:

- Home Medications
- Allergies and Reactions
- Medications Administered in ED
- ED Provider Note
- Mental Status and Orientation Assessment
- Reason for Transfer and Plan of Care
- Tests and/or Procedures Performed
- Tests and/or Procedures Results

For ALL data elements, the definition of 'sent' includes the following documentation:

- Hard copy sent directly with the patient, or
- Communicated via fax or phone within 60 minutes of patient departure, or
- Immediately available via shared Electronic Health Record (EHR) or Health Information Exchange (HIE) (see definition below)

For purposes of this measure, a shared electronic health record (EHR) is defined as one where data entered into the system is immediately available at the receiving site. Facilities using the same EHR vendor or a Health Information Exchange (HIE) cannot assume immediate access by the receiving facility to the transferred patient's records.

ED Transfer Communication (EDTC) Initial Patient Population

The population of the EDTC measure is defined by identifying those patients admitted to the emergency department who were then discharged, transferred, or returned to these facilities:

Inclusions:

- Acute Care Facility – Cancer Hospital or Children's Hospital – Including emergency department
- Acute Care Facility – Critical Access Hospital – Including emergency department
- Acute Care Facility – Department of Defense or Veteran's Administration – Including emergency department
- Acute Care Facility- General Inpatient Care – Including emergency department
- Hospice – healthcare facility
- Other health care facility*, including discharge, transfer or return to:
  - Extended or Intermediate Care Facility (ECF/ICF)
  - Long Term Acute Care Hospital (LTACH)
  - Long Term Care Facility
  - Nursing Home or Facility, including Veteran's Administration Nursing Facility
  - Psychiatric Hospital or Psychiatric Unit of a Hospital
  - Rehabilitation Facility, including Inpatient Rehabilitation Facility/Hospital or Rehabilitation Unit of a Hospital
  - Skilled Nursing Facility (SNF), Sub-Acute Care, or Swing Bed
  - Transitional Care Unit (TCU)

*Other health care facilities MUST be included in the population.

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

Denominator Statement: Transfers from an ED to another healthcare facility

Included Population: All transfers from an ED to another healthcare facility

Excluded Populations: Patients observation status.

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).

ED Transfer Communication (EDTC) Initial Patient Population

The population of the EDTC measure is defined by identifying those patients admitted to the emergency department who were then discharged, transferred, or returned to these facilities:

Inclusions:

- Acute Care Facility – Cancer Hospital or Children's Hospital – Including emergency department
- Acute Care Facility – Critical Access Hospital – Including emergency department

- Acute Care Facility – Department of Defense or Veteran's Administration – Including emergency department
- Acute Care Facility- General Inpatient Care – Including emergency department
- Hospice – healthcare facility
- Other health care facility*, including discharge, transfer or return to:
  - Extended or Intermediate Care Facility (ECF/ICF)
  - Long Term Acute Care Hospital (LTACH)
  - Long Term Care Facility
  - Nursing Home or Facility, including Veteran's Administration Nursing Facility
  - Psychiatric Hospital or Psychiatric Unit of a Hospital
  - Rehabilitation Facility, including Inpatient Rehabilitation Facility/Hospital or Rehabilitation Unit of a Hospital
  - Skilled Nursing Facility (SNF), Sub-Acute Care, or Swing Bed
  - Transitional Care Unit (TCU)

*Other health care facilities MUST be included in the population.

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

All emergency department patients not discharged to another healthcare facility.

Those admitted, sent home, left AMA, those on observations status, etc.

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.)*

Exclusions:

- AMA (left against medical advice)
- Expired
- Home, including:
  - Assisted Living Facilities
  - Board and care, foster or residential care, group or personal care homes, and homeless shelters
  - Court/Law Enforcement – includes detention facilities, jails, and prison
  - Home with Home Health Services
  - Outpatient Services including outpatient procedures at another hospital, Outpatient Chemical Dependency Programs, and Partial Hospitalization
- Hospice-home
- Not documented/unable to determine
- Observation Status

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.)*

Stratification by discharge destination is suggested to target groups where improvement is needed specifically separating acute care vs long term care transfers.

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:

Continuous variable, e.g. average

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.)*

The measure is reported as an average of the patient observations scores from the facility. The individual patient's score is the sum of the elements scores which use an all-or-none approach. If an individual patient's record is missing one element then the patient's score is zero. Reporting of element level scores may be useful for improvement or reporting. Data elements are identified for the measure. If the data element is not appropriate for the patient, items scored as NA (not applicable or not available) are counted in the element as a positive, or 'yes,' response and the patient will meet the element criteria. The patient will either need to meet the criteria for all of the data elements (or have an NA) to pass the element. Measurement stratification by discharge destination may help identify areas for targeted improvement activities.

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.

Hospitals need to submit a minimum of 45 cases per quarter from the required population.

A hospital may choose to sample and submit more than 45 cases. Hospitals that choose to sample have the option of sampling quarterly or sampling monthly. Hospitals whose initial patient population size is less than the minimum number of 45 cases per quarter for the measure cannot sample and should submit all cases for the quarter.

Hospital samples must be monitored to ensure that sampling procedures consistently produce statistically valid and useful data. Sample cases should be randomly selected in such a way that the individual cases in the population have an equal chance of being selected.

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.

n/a

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

If other, please describe in S.18.

Claims, Electronic Health Data, Electronic Health Records, Management Data, Paper Medical Records

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

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

Hospitals submit information to their State Flex Coordinators and the State Flex Coordinators submit info to the MBQIP data bank.

Hospitals use a variety of tools to collect information from their ED records including paper records, electronic records and external vendors.

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)*

Available in attached appendix at A.1

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

Facility

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

Emergency Department and Services, 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.)*

2. Validity – See attached Measure Testing Submission Form

NQF_testing_attachment_11.6.19-637087248040991410__12.18.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.

Yes

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.

Yes

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.

No - This measure is not risk-adjusted

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

Measure Number (*if previously endorsed*)**:** 0291
Measure Title: Emergency Transfer Communication Measure
Date of Submission: Fall 2019

Type of Measure:

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

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:	☐

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*).

The Federal Office of Rural Health Policy (FORHP) has set up a reporting process for the Emergency Department Transfer Communication Measures (EDTC) Critical Access Hospitals (CAH): each CAH provides data to the State Flex Office or the State office of Rural Health or Primary Health, which is then compiled into an Excel template supplied by FORHP. The raw Excel data file from each state is submitted to FORHP, which subsequently submits that data to Telligen. Telligen then generates state and hospital reports, which are distributed

back to State Flex Offices via FORHP Project Officers. State Flex Coordinators and critical access hospitals (CAHs) utilize the EDTC reports to implement quality improvement initiatives.

The data used to generate Cohen's Kappa to test the reliability of element level measurement was from a June 2019 pilot test of the 2019 specifications from 34 CAHs and Stratis Health.

IRR data from previous use is included here. The IRR activities includes previous versions of specifications. This from 2009 field test data includes the current data elements.

The Telligen database is also the source of the data on the measures used for comparison in the validity section of this document and the reliability testing of the composite measure.

1.3. What are the dates of the data used in testing? Quarter 4 2017, June 2019

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: Click here to describe	☐ other: Click here to describe

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*)

No information is available about the source of the data, i.e., whether it is abstracted from paper or electronic records. The data is submitted to the central data base in an excel format. CAH hospitals are not required to identify if the abstraction is from paper or electronic records.

All records that were available were used. The analysis was done on a sample of submissions but using all of the submissions available for these measures in the time frame.

For Reliability testing 1348 CAH hospitals' data was used in the Beta-Binomial calculations. This included 190650 patient charts that are aggregated to the hospital level. Only the 8 elements from the new specifications are included in these analyses.

For IRR reliability testing 34 hospitals with approximately 83 patient chart records were used. Twenty-one (of 104) charts were excluded because the submission criteria were not followed. Raw agreement between the expert reviewer and hospital reviewers is also presented.

For the first quarter of 2009 data, 197 records were abstracted at 23 hospitals in three states. In 134 or 68% of those records the hospital abstractors findings agreed 100% with the QIO staff abstraction. In the second quarter, 165 charts were abstracted at 19 hospitals. 136, 82.4%, of those records the hospitals abstraction findings agreed 100% with the QIO staff abstraction. The number of inconsistencies in abstraction decreased from 74 to 29 from the first quarter to the second quarter.

For Validity testing 110 entities were included in the Fibrinolytics measure, 200 entities were included in the Aspirin measure and 907 entities were included in the Immunization measure for analysis. The number was based on the number of hospitals who had submitted data for the comparison measures.

Number of entities	**EDTC**
Immunization	907
Administration of Fibrinolytics	110
Administration of Aspirin	200

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*)

All patients who were transferred from the emergency department to another facility that provided onsite care provided by trained medical professionals are included. Diagnosis, age race, gender, etc., were not considered in the inclusion criteria.

For reliability Beta Binomial calculations approximately 190650 patients were included in the data.

For reliability/IRR calculations approximately 83 patients were included in the data.

For reliability/IRR calculations using previously reported 2009 data 197 records were abstracted at 23 hospitals in three states.

For Validity testing – the EDTC sample compared to the comparison measures IMM had 35777 patients, OP4 had 7872 patients, and OP2 had 4527 patients.

Number of patient charts	EDTC elements
Immunization	35777
Administration of Fibrinolytics	4527
Administration of Aspririn	7872

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.

For reliability Beta Binomial calculations approximately 190650 patients were included from Telligen data.

For reliability/IRR calculations approximately 83 patients were included from the June 2019 Pilot IRR data.

For reliability/IRR calculations using previously reported 2009 data 197 records were abstracted at 23 hospitals in three states.

For Validity testing – the EDTC sample compared to the comparison measures IMM had 35777 patients, OP4 had 7872 patients, and OP2 had 4527 patients.

Number of patient charts	EDTC elements
Immunization	35777
Administration of Fibrinolytics	4527
Administration of Aspirin	7872

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.

We have no reason to suggest that a hospital discharge or transfer processes would be influenced by social factors of the individual patients.

No patient level data was used.

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*)

We conducted reliability testing at the performance score level using a Beta-Binomial model and on the individual element level using Kappa score from Inter Rater Reliability (IRR).

All analysis and data cleaning were completed using SAS 9.4 in SAS Studio 3.8.

Performance score level analysis was estimated using Adams methods as in https://www.rand.org/content/dam/rand/pubs/technical_reports/2009/RAND_TR653.pdf.

The macro for calculating the alpha and beta values for our beta-binomial distribution was developed by:

Ian Wakeling - Qi Statistics (email: ian@qistatistics.co.uk) Web site: www.qistatistics.co.uk

Reference: http://www.qistats.co.uk/BetaBinomial.html

In this report, we estimate reliability with a beta-binomial model. The beta-binomial is a natural model for estimating the reliability of simple pass/fail rate measures. There are also computational advantages to using the beta-binomial model, which is based on the beta distribution for the "true" hospital scores. The beta distribution is a very flexible distribution on the interval from 0 to 1. The beta-binomial model assumes the hospital's score is a binomial random variable conditional on the hospital's true value that comes from the beta distribution.

This method underscores that reliability is not just a property of a measure set but also depends what population is used to estimate the reliability. Whether a set of measures is useful for profiling hospitals depends on how different the hospitals are from one another. Measures that may be useful in one group of hospitals may not be useful in another group with little hospital to hospital variation. Similarly, as the hospitals under study increase their performance, the reliability may decrease if the hospital to hospital variance decreases over time. This is especially true as measures hit the upper limits of their ranges.* (*The Reliability of Provider Profiling A Tutorial John L. Adams Prepared for the National Committee for Quality Assurance HEALTH)

IRR Methods were used to assess the reliability at the data element level.

IRR was calculated using SAS.

We reviewed 8 elements for 83 hospital records. They were first reviewed by the hospital and then reviewed by an expert. To evaluate the agreement between the hospital and the expert, we used Cohen's Kappa coefficients as an estimation of agreement. Cohen's Kappa is widely known as an appropriate measure of agreement when looking at binary (pass/fail) agreement between two reviewers. A Kappa coefficient of 0 indicates that there was no negative or positive correlation between the two raters. Alternatively, a value of negative 1 indicates complete disagreement while positive 1 indicates complete agreement. There is no official rating of agreement in terms of strength but Kappa coefficients above 0.75 can be seen as substantial positive agreement[1].

The raw agreement between the expert reviewer and the hospital reviewers showed high level of agreement with opportunity for clarification. Please note these are with the new specification revisions. Additional training will be provided Q4 2019.

Element_1	Element_2	Element_3	Element_4	Element_5	Element_6	Element_7	Element_8
73.1%	69.2%	73.1%	72.1%	70.2%	81.7%	75.0%	75.0%

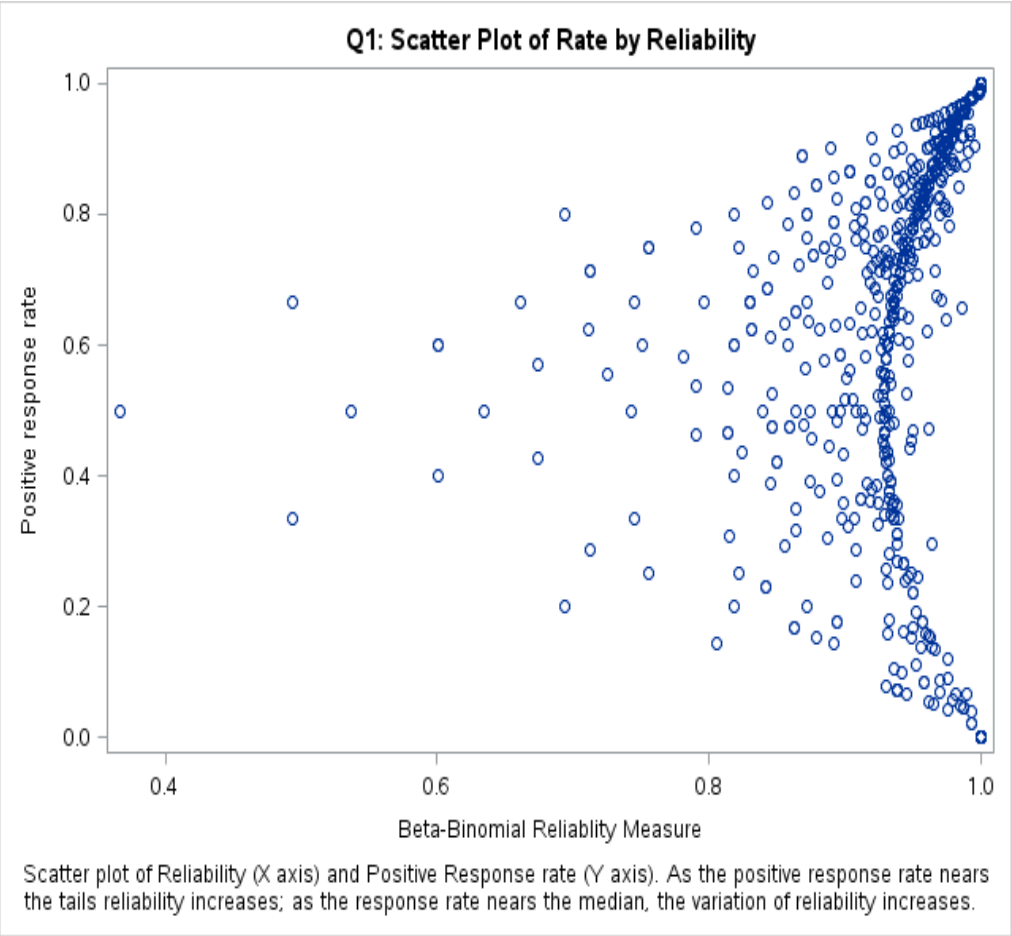
For the first quarter of 2009 data, 197 records were abstracted at 23 hospitals in three states. In 134 or 68% of those records the hospital abstractors findings agreed 100% with the QIO staff abstraction. In the second quarter, 165 charts were abstracted at 19 hospitals. 136, 82.4%, of those records the hospitals abstraction findings agreed 100% with the QIO staff abstraction. The number of inconsistencies in abstraction decreased from 74 to 29 from the first quarter to the second quarter.

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*)

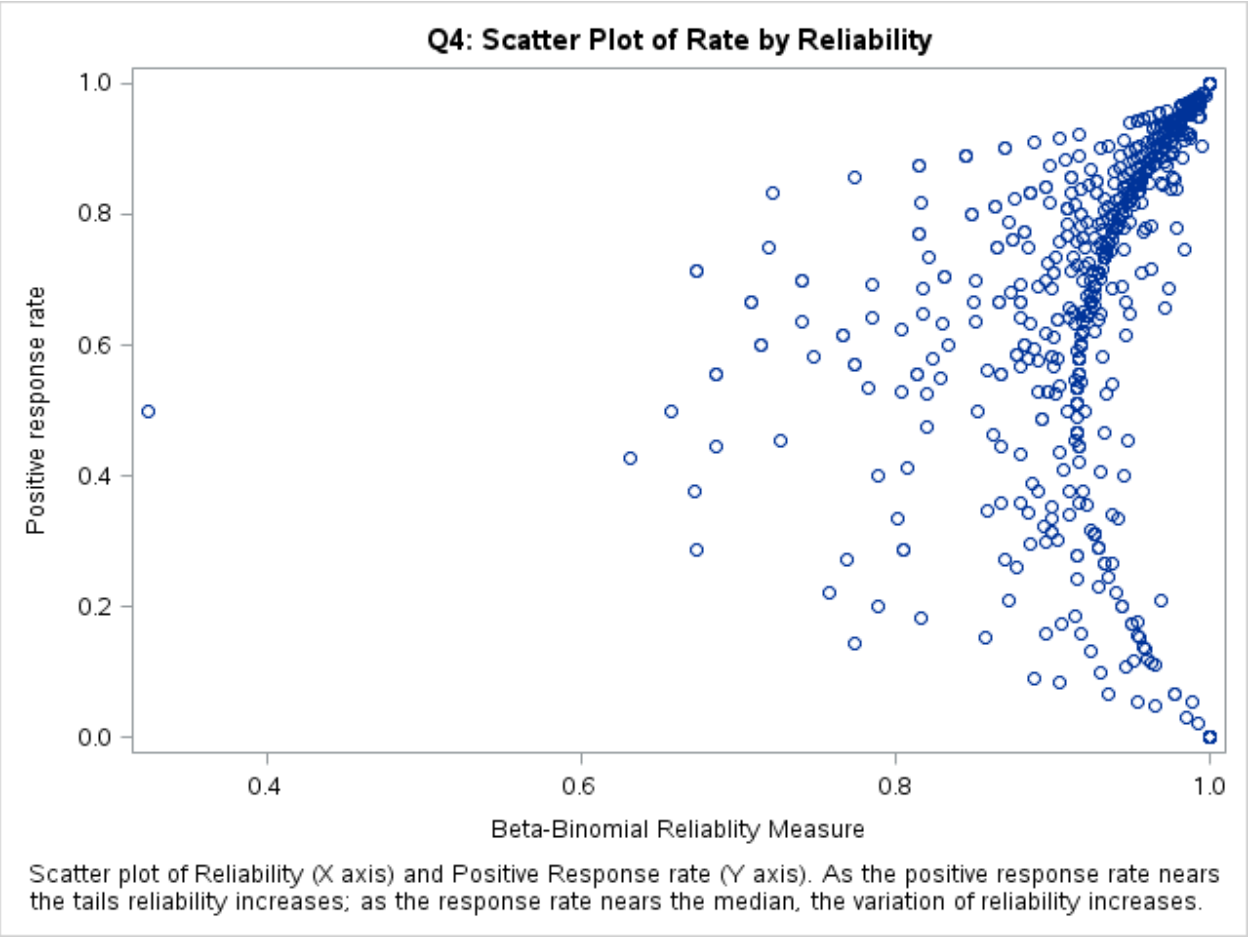
Performance measure score reliability was tested using a Beta-Binomial distribution.

Rate and Reliability Distributions – (n: 1,185)
Reliability from Beta-Binomial Distribution (alpha = 1.4651; beta = 0.4188)

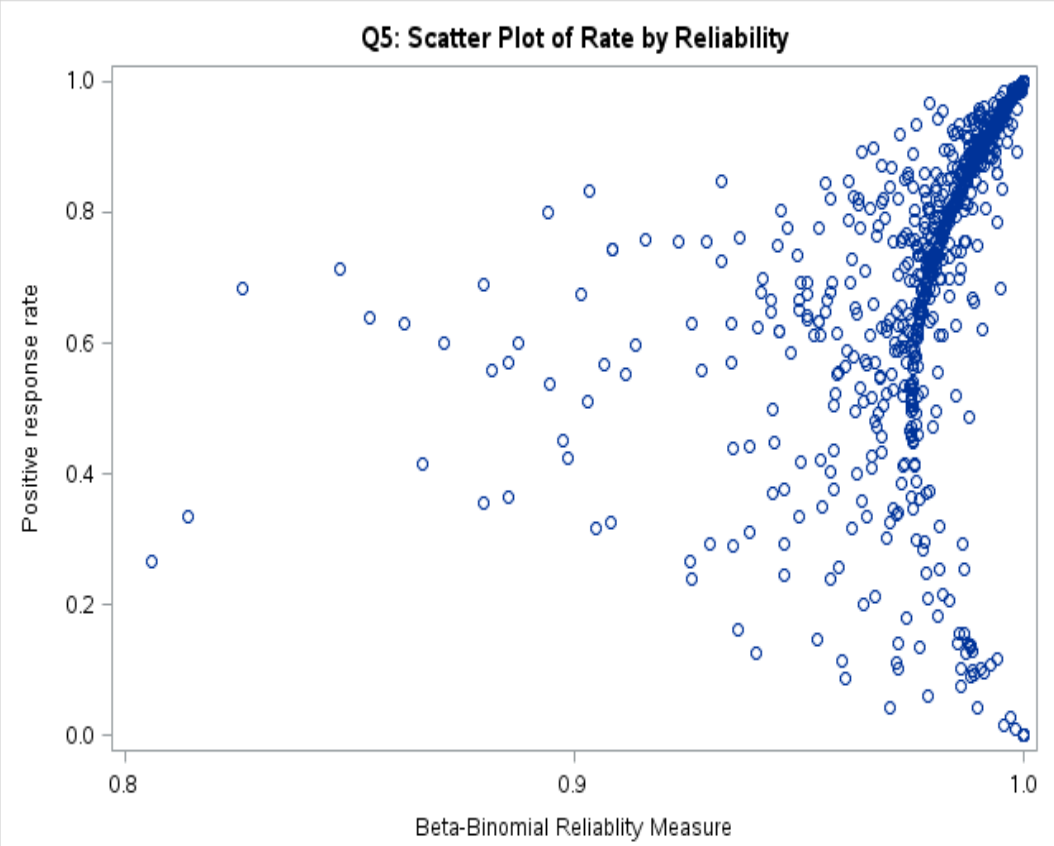
Variable	**Mean**	**Std Dev**	**Min.**	**10th %**	**25th %**	**50th %**	**75th %**	**90th %**	**95th %**	**Max.**
Q1 Score (range: 0-1)	0.786	0.253	0	0.407	0.689	0.878	0.978	1.0	1.0	1.0
Q1 – Reliability (range: 0-1)	0.980	0.060	0.324	0.890	0.926	0.965	1.0	1.0	1.0	1.0



Scatter plot of Reliability (X axis) and Positive Response rate (Y axis). As the positive response rate nears the tails reliability increases; as the response rate nears the median, the variation of reliability increases.



Scatter plot of Reliability (X axis) and Positive Response rate (Y axis). As the positive response rate nears the tails reliability increases; as the response rate nears the median, the variation of reliability increases.



Scatter plot of Reliability (X axis) and Positive Response rate (Y axis). As the positive response rate nears the tails reliability increases; as the response rate nears the median, the variation of reliability increases.

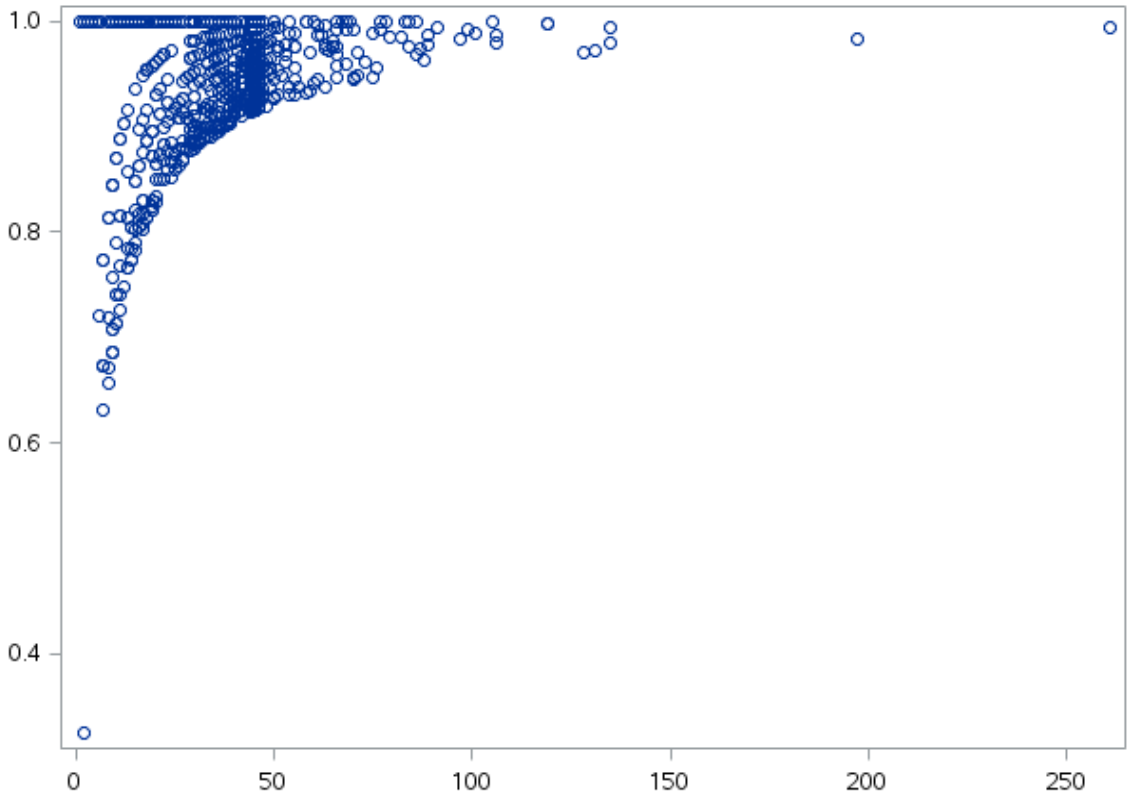
The raw agreement between the expert reviewer and the hospital reviewers showed high level of agreement with opportunity for clarification. Please note these are with the new specification revisions. Additional training will be provided Q4 2019.

Element_1	Element_2	Element_3	Element_4	Element_5	Element_6	Element_7	Element_8
73.1%	69.2%	73.1%	72.1%	70.2%	81.7%	75.0%	75.0%

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?*)

- Interpretation: Based on the observed average reliability of 0.95 in both quarters 1 and 4, and the patterns of variation observed on the visualizations above, we can say that these measures were reliable. A key observation is that, for hospitals with the highest scores, the reliability of the measure increases, the same can be said for the hospitals with the lowest scores. Most of the variation in reliability occurs in the hospitals with a success rate between 0.2 & 0.6, though in both quarters 75% of the hospitals had a success rate above 65%. More analysis should be done into the drivers of reliability and inter hospital variation. As a visual analysis, below is a plot of reliability by number of trials for each hospital in quarter 4.
-

- Scatter plot of Reliability (Y) axis and number of responses (X) axis: As the response rates increase reliability increases.



- Scatter plot of Reliability (Y) axis and number of responses (X) axis: As the response rates increase reliability increases.

IRR results

There are several operational definitions of "inter-rater reliability", reflecting different viewpoints about what is a reliable agreement between raters.[1] The operational definitions of agreement: Reliable raters agree with the "official" rating of a performance.

We tested the agreement between the independent hospital based raters with the 'official' rating (Stratis Health, Robyn Carlson) of abstraction.

The raw agreement between the expert reviewer and the hospital reviewers showed high level of agreement with opportunity for clarification. Please note these are with the new specification revisions. Additional training will be provided Q4 2019.

Element_1	Element_2	Element_3	Element_4	Element_5	Element_6	Element_7	Element_8
73.1%	69.2%	73.1%	72.1%	70.2%	81.7%	75.0%	75.0%

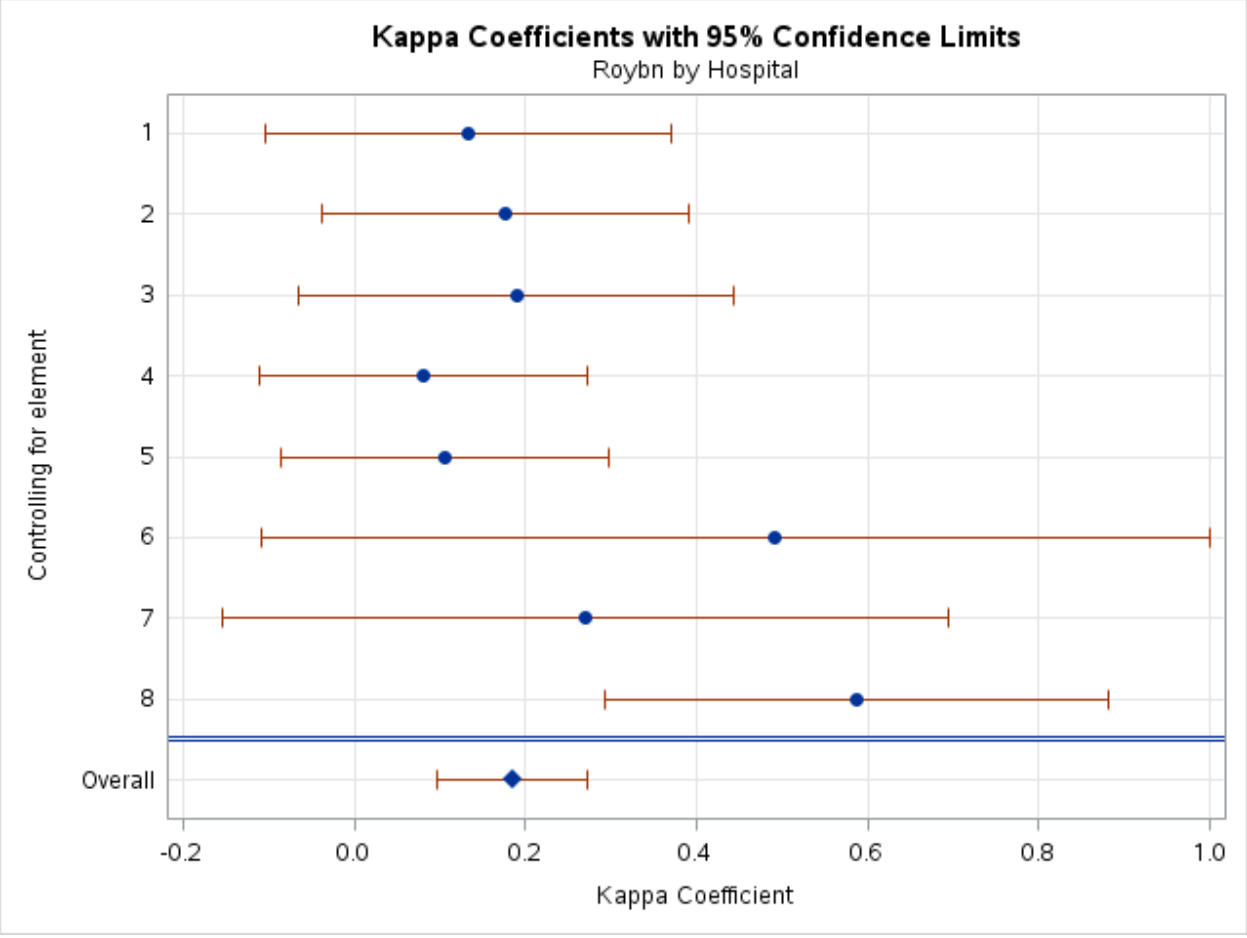
Overall, the hospital and expert had slight to fair crude non-zero agreement (Kappa = 0.224; 95% CI: [0.122 - 0.327]). The individual elements had a range of Kappa values, element 4 had the lowest agreement (Kappa=0.08) while element 5 had the highest (Kappa=0.59). The Element adjusted overall agreement within our sample was slightly lower than our observed

crude value (Kappa = 0.185; 95% CI: [0.097 - 0.273]). Figure 1 below shows the element specific and overall Kappa coefficients along with a. 95% confidence interval. An overall chi-square test indicated that there was statistically significant difference between the individual element's agreement values (p=0.171).

Our analysis found that both adjusted/unadjusted agreement between the hospital and expert raters was slight to fair. There was also no evidence that a single element of the 8 was responsible for this result as the kappa coefficients were homogeneous, though the elements with the lowest agreement were 4, 5, & 1 respectively.

Previous IRR analysis showed good agreement. The specifications with respect to the measures have changed little. The difference between the two may be due to the availability of resources. The 2019 pilot of specifications purposely did not provide support during the data collection period to test the use of the manual. Clearly the additional on call support is necessary for ongoing reliability.

For the first quarter of 2009 data, 197 records were abstracted at 23 hospitals in three states. In 134 or 68% of those records the hospital abstractors findings agreed 100% with the QIO staff abstraction. In the second quarter, 165 charts were abstracted at 19 hospitals. 136, 82.4%, of those records the hospitals abstraction findings agreed 100% with the QIO staff abstraction. The number of inconsistencies in abstraction decreased from 74 to 29 from the first quarter to the second quarter.



References:

[1]Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. Fam med, 37(5), 360-363.

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 <u>performance measure score</u> 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)*

Pearson correlation coefficients were calculated comparing three ED quality performance measures that are related to process of care to the EDTC measures.

Three ED measures were used to compare with groupings of the 8 EDTC elements included in the new 2019 EDTC measurement. EDTC data is not available for the each element but subsections 4, 5, and 7 represent the elements that remain in the revised measure.

2019 EDTC Elements by subsection are

- Subsection 4
  - Medications Administered in ED
  - Allergies and Reactions
  - Home Medications
- Subsection 5
  - ED Provider Note
  - Mental Status and Orientation Assessment
  - Reason for Transfer and Plan of Care
- Subsection 7
  - Tests and Procedures Performed
  - Tests and Procedures Results

The comparison measures are IMM2 (Immunization rate of Pts), OP 2 (Fibrinolytic therapy received in the ER within 30 minutes), and OP 4 (Aspirin upon ER arrival).

Analysis was done using a one-tailed model.

The sample size for IMM2 is the highest and most amenable to comparison. Sample size for OP 4 is acceptable. Because CAH hospitals rarely provide fibrinolytic care the sample sizes are limited for OP 2.

Data for this testing was from Q4, 2017

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

Pearson Correlation Coefficients were generated comparing three Emergency Department Process measures to the EDTC element combinations.

The Comparison Measures are:

IMM2 -Immunization screening for influenza vaccine status and vaccinated prior to transfer

OP2 - Fibrinolytic Therapy Received Within 30 Minutes of ED arrival

OP4 - Aspirin therapy within 24 hours before ED arrival or prior to transfer

These comparison measures were chosen because they represent 0-1 possible outcomes and because they are process measures for patients in the same setting the ED.

The EDTC data that is used in this analysis is separate into 3 groups of elements. This is done because our data is from the previous EDTC specifications. The new specifications only include the elements in these 3 subsets. The data we have cannot be separated into the 8 individual elements nor can they be aggregated into one score.

The quality constructs for these comparisons apply to the same population (Emergency Department patients) at the same care location (Emergency departments), at the same type of facilities (CAHs), from the same data sources, and in the same time frame. These quality measures are all process measures with 0-1 results.

These quality constructs expressed by these EDTC measures vary from the comparison measures because of the nature of the processes and the facility structures that support these measures. We did not expect high correlation scores but did expect positive and statistically significant correlation.

As expected the results showed positive and significant correlation between IMM, OP4 and EDTC 4, 5, 7.

Correlations – One Tailed – Q4

In the table below the Pearson Correlation coefficients for the Immunization Rate and ER Communication measures in Q4 are presented. The Immunization Rate of patients in the ER is positively and significantly correlated to all three of the Emergency Department Transfer Communication measure, at the 1% level of significance. The implication is that increased levels of communication in the ER are positively related to higher levels of immunization in older patients.

CAHs with attention to ED quality of care follow standards of care for immunizations and communication processes. Both administration of proper influenza immunization and communication of appropriate information signal an attention to future health and ongoing care for these similar populations.

Correlations- Q4 Immunization and EDTC 4, 5, 7

		imm2_ratio_Q4	Q4_EDTC4	Q4_EDTC5	Q4_EDTC7
imm2_ratio_Q4	Pearson Correlation	1	.236**	.255**	.218**
	Sig. (1-tailed)		.000	.000	.000
	N	984	907	907	907
Q4_EDTC4	Pearson Correlation	.236**	1	.971**	.976**
	Sig. (1-tailed)	.000		.000	.000
	N	907	1185	1185	1185
Q4_EDTC5	Pearson Correlation	.255**	.971**	1	.980**
	Sig. (1-tailed)	.000	.000		.000
	N	907	1185	1185	1185
Q4_EDTC7	Pearson Correlation	.218**	.976**	.980**	1
	Sig. (1-tailed)	.000	.000	.000	
	N	907	1185	1185	1185

**. Correlation is significant at the 0.01 level (1-tailed).

In the table below the Pearson Correlation coefficients for Fibrinolytic therapy and ER Communication measures in Q4 are presented. While the Fibrinolytic therapy received in the ER within 30 minutes is positively correlated to all three of the Emergency Department Transfer Communication measure, the values are not significant. There is no statistical support for the expectation that increased levels of communication in the ER are positively related to rapid Fibrinolytic therapy delivery in the ER.

Few CAHs provide urgent ER based fibrinolytic treatment for chestpain/AMI patients prior to transfer, therefore this sample is very small.

CAHs with attention to ED quality of care follow standards of care for fibrinolytic and communication processes. Both administration of proper fibrinolytics and communication of appropriate information signal an attention to future health and ongoing care for these similar populations.

Correlations – Q4 Fibrinolytic Administration and EDTC 4, 5, 7

	op2_ratio_Q4	Q4_EDTC4	Q4_EDTC5	Q4_EDTC7

op2_ratio_Q4	Pearson Correlation	1	.055	.076	.062
	Sig. (1-tailed)		.286	.214	.261
	N	119	110	110	110
Q4_EDTC4	Pearson Correlation	.055	1	.971**	.976**
	Sig. (1-tailed)	.286		.000	.000
	N	110	1185	1185	1185
Q4_EDTC5	Pearson Correlation	.076	.971**	1	.980**
	Sig. (1-tailed)	.214	.000		.000
	N	110	1185	1185	1185
Q4_EDTC7	Pearson Correlation	.062	.976**	.980**	1
	Sig. (1-tailed)	.261	.000	.000	
	N	110	1185	1185	1185

**. Correlation is significant at the 0.01 level (1-tailed).

In the table above the Pearson Correlation coefficients for provision of Aspirin and ER Communication measures in Q4 are presented. The Aspirin delivery rate upon ER arrival is positively and significantly correlated to all three of the Emergency Department Transfer Communication measure, at the 5% level of significance. The implication is that increased levels of communication in the ER are positively related to higher levels of Aspirin delivery upon arrival to the ER.

CAHs with attention to ED quality of care follow standards of care for Aspirin administration for patients with chest pain and communication processes. Both administration of proper Aspirin and communication of appropriate information signal an attention to future health and ongoing care for these similar populations.

Correlations Q4 Aspirin administration and EDTC 4, 5, 7

		op4_ratio_Q4	Q4_EDTC4	Q4_EDTC5	Q4_EDTC7
op4_ratio_Q4	Pearson Correlation	1	.174**	.165**	.164*
	Sig. (1-tailed)		.007	.010	.010
	N	215	200	200	200

Q4_EDTC4	Pearson Correlation	.174**	1	.971**	.976**
	Sig. (1-tailed)	.007		.000	.000
	N	200	1185	1185	1185
Q4_EDTC5	Pearson Correlation	.165**	.971**	1	.980**
	Sig. (1-tailed)	.010	.000		.000
	N	200	1185	1185	1185
Q4_EDTC7	Pearson Correlation	.164*	.976**	.980**	1
	Sig. (1-tailed)	.010	.000	.000	
	N	200	1185	1185	1185

**. Correlation is significant at the 0.01 level (1-tailed).

*. Correlation is significant at the 0.05 level (1-tailed).

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?*)

CAHs with attention to ED quality of care follow standards of care for Influenza Immunizations screening and administration, Aspirin administration for patients with chest pain, and communication processes. Both administration of proper administration of Influenza Immunizations, Aspirin for those with chest pain, and communication of appropriate information signal an attention to future health and ongoing care for these similar populations. The quality constructs vary due to attention to the process of care delivery vs the process of communication. We expect a weak but positive and significant relationship.

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section 2b4*

All patients transferred from an ED to another skilled care providing facility are eligible for this measure.

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

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across*

measured entities, and impact on performance measure scores)

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*)

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 2b5.

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

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

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.

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.

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?

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

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

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

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.**

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

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

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

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

2b3.9. Results of Risk Stratification Analysis:

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*)

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*)

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)*

Data from a CMS directed a special innovation project lead by Stratis Health provided data from 100 CAH in 8 states. This data provided insight into score differences between discharge destinations.

Tellegin data was used to generate description of differences across the entities with simple descriptive statistics. A sample of 1185 CAH reported EDTC 4, 5 and 7 and are reported here.

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*)

The EDTC measures have been developed and used since 2006. EDTC improvements have been supported by HRSA, FORHP, Flex program and QIOs. Over time improvements have been demonstrated overall. Hospitals that are actively improving EDTC communication have shared process support tools, hospital transfer agreements, and other communication process mechanisms.

Once the process has been restructured with training and tools EDTC hospital to hospital scores often reach 100%. Hospital to non-acute facilities communication has not improved at the same rate or to the 100% scoring.

Stratis Health lead a one-year Centers for Medicare & Medicaid (CMS) special innovation project to assist eight Medicare Quality Improvement Organizations (QIOs), in Iowa, Maine, Missouri, Nebraska, Oklahoma, West Virginia, Wisconsin, and Wyoming, to train CAHs to collect and report seven composite ED transfer communication (EDTC) measures, identify gaps and opportunities for improvement, and begin planning to improve the transfer communication process and results.

The EDTC measures were originally developed by Stratis Health and the University of Minnesota Rural Health Research Center and endorsed by the National Quality Forum in 2007. Nearly 100 critical access hospitals (CAHs) across eight states worked on using measures to evaluate communication for transitions of care during emergency department (ED) transfers. Participating CAHs abstracted medical records to collect data on the EDTC measures. CAHs submitted data through their QIOs to Stratis Health for benchmarking with other participating facilities.

Improved process measures—56% relative improvement rate

Participating CAHs increased their percentage of medical records meeting all of the EDTC data elements over the course of the project from 28.26 to 44.13 percent—for a relative improvement rate of 56 percent.

The hospitals used the results to develop and implement improvements focused on better documentation and communication processes.

Analysis of different receiving facilities, and different levels of care data transfer analysis showed that CAH performance on the measures varied with where patients were transferred.

The percentage of medical records containing all necessary patient data transferred in a timely manner was 36.79 percent for acute care hospital transfers, but only 20.19 percent for transfers to other health care facilities, such as nursing homes.

Medical Records With All EDTC Patient Data

CAH Transfers To	Percentage Complete
Acute care hospital	36.7
Other health care facilities	20.19

This data highlights an opportunity for improved transition communication from EDs to long term care facilities, by working with local nursing homes to develop standard communication and transition processes.

The EDTC measure is included in phase three of the Medicare Beneficiary Quality Improvement Project (MBQIP). Starting fall 2014, CAHs nationwide can collect and submit the EDTC measures. MBQIP is a program of the Health Resources and Services Administration (HRSA) funded Office of Rural Health Policy's (ORHP) Medicare Rural Hospital Flexibility Program (Flex).

Data from Telligen does not identify destination therefore the significance difference between entities for acute and non-acute transfers cannot be generated.

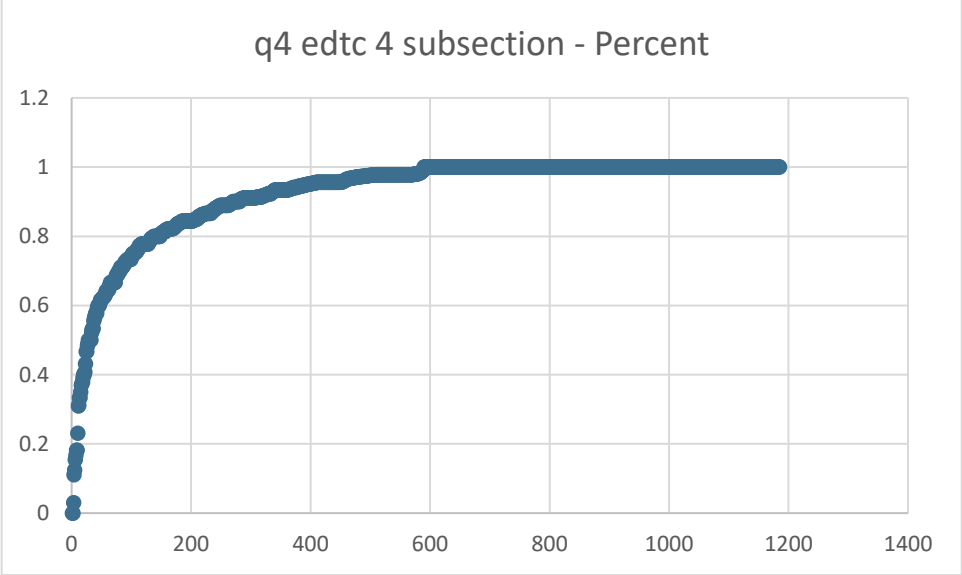
Below are descriptive statistics of the distribution of Quarter 4 EDTC 4, EDTC 5, EDTC 7 from the Telligen data. This data does not differentiate between acute and non-acute transfers.

q4 edtc 4 %	
Mean	0.93
Standard Error	0.00
Median	1.00
Mode	1.00
Standard Deviation	0.14
Sample Variance	0.02
Kurtosis	11.91
Skewness	-3.10
Range	1.00
Minimum	0.00
Maximum	1.00
Sum	1096.39

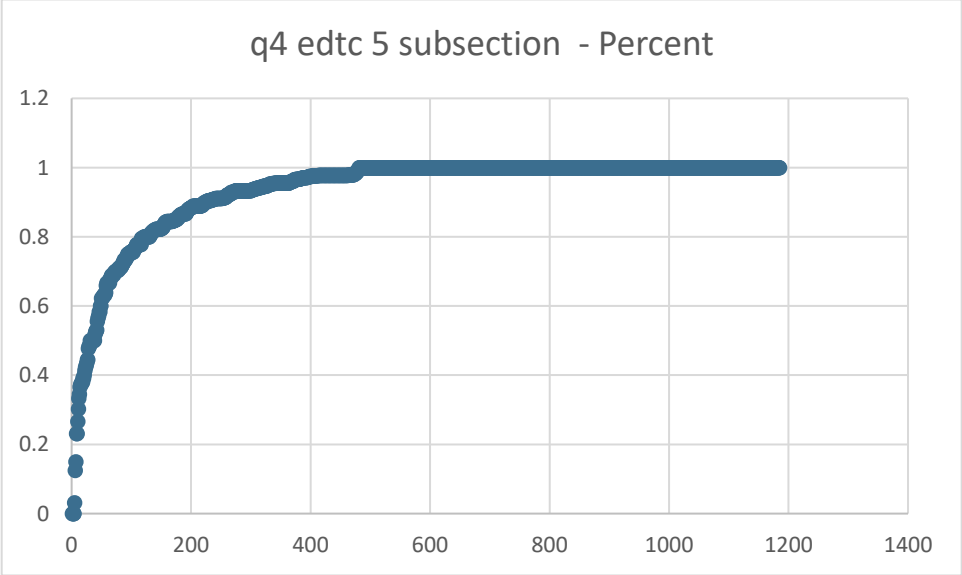
Count	1185.00

q4 edtc 5 %	
Mean	0.93
Standard Error	0.00
Median	1.00
Mode	1.00
Standard Deviation	0.14
Sample Variance	0.02
Kurtosis	13.71
Skewness	-3.38
Range	1.00
Minimum	0.00
Maximum	1.00
Sum	1107.68
Count	1185.00

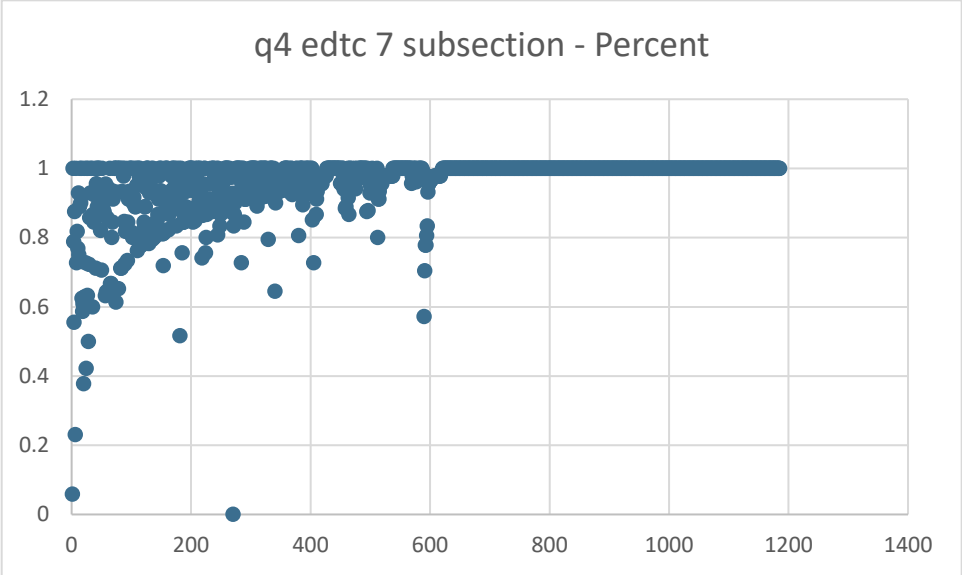
q4 edtc 7 %	
Mean	0.96
Standard Error	0.00
Median	1.00
Mode	1.00
Standard Deviation	0.09
Sample Variance	0.01
Kurtosis	33.26
Skewness	-4.73
Range	1.00
Minimum	0.00
Maximum	1.00
Sum	1142.43
Count	1185.00



Scatter plot Distribution of Measure Scores for EDTC subsection 4 (Y) axis and Observations (X) axis. Display shows a mean of .93 and a SD of 0.14



Scatter plot Distribution of Measure subsection 5 Scores (Y) axis and observations (X) axis. Display shows a mean of .93 and a SD of 0.14



Scatter plot Distribution of Measure Subsection 7 Scores (Y) axis and observations (X) axis. Display shows a mean of .96 and a SD of 0.09.

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?*)

The data include transfers to all facilities that provide clinically trained staff. CAH to tertiary hospital communication has improved over time. CAH transfers to non-acute facilities such as nursing homes, assisted living, detox centers still needs improvement. That data assessment is not differentiated in this data set. Changes in submission expectations are being considered to help identify these improvement opportunities.

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

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

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not***

demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.

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

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*)

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*)

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

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

The EDTC measure looks for completeness of data elements in a communication between two facilities. Its purpose is to identify the information that is missing from or not identified and/or included in the communication.

Because CAHs by definition limit stays to 72 hours and have a structure/process developed with supporting hospitals to transfer patients with higher acuity or special care needs beyond what is available in the CAH; this measure attempts to measure one of CAHs' key roles, that of a communicator of initial information to promote continuity of care and decrease redundant testing.

We examine the charts from the sending hospital as we evaluate their performance as a transferring hospital. To assess the sensitivity or specificity of missing data abstraction we would need to look at the receiving facilities records to identify cases where records were identified as sent but were not sent; or records that were identified as not sent but were in fact sent. We do not have access to the data at the receiving hospital. Without access to the receiving hospital data we are unable identify missing data nor to estimate the errors related to that.

The measure itself looks at data that might not be sent. Analysis of potential data abstraction errors was presented in the IRR Kappa statistics.

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*)

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*)

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.

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

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

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

ALL data elements are in defined fields in electronic health records (EHRs)

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).

Hospitals with electronic medical records shared with the patient´s destination can capture these elements easily. Hospitals who do not share electronic medical records with the destination hospital require additional paper trails for capture of necessary process elements.

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.

The measure was originally tested in 2003. Since that time we have expanded the application to include transfers to non-acute settings. We continue to clarify the specifications.

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)*.

none.

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)
Payment Program Regulatory and Accreditation Programs	Public Reporting Program Name: Minnesota Statewide Quality Reporting and Measurement System. http://www.mnhealthscores.org/hospital-quality-patient-experience-ratings Quality Improvement (external benchmarking to organizations) MN statewide forms http://www.health.state.mn.us/healthreform/measurement/index.html Medicare Beneficiary Quality Improvement Project (MBQIP) https://www.ruralcenter.org/tasc/mbqip Quality Improvement (Internal to the specific organization) Medicare Beneficiary Quality Improvement Project (MBQIP) https://www.ruralcenter.org/tasc/mbqip

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

MBQIP:

• Name of program and sponsor: The Medicare Beneficiary Quality Improvement Project (MBQIP) is a quality improvement activity under the Medicare Rural Hospital Flexibility (Flex) grant program of the Health Resources and Services Administration's Federal Office of Rural Health Policy (FORHP).

• Purpose: The goal of MBQIP is to improve the quality of care provided in critical access hospitals (CAHs), by increasing quality data reporting by CAHs and then driving quality improvement activities based on the data. This project provides an opportunity for individual hospitals to look at their own data, measure their outcomes against other CAHs and partner with other hospitals in the state around quality improvement initiatives to improve outcomes and provide the highest quality care to each and every one of their patients. In order to access the full scope of Flex resources available to them, CAHs must be reporting a minimum threshold of quality measures. Reporting of EDTC is one of the measures that is included in that criteria. The reporting threshold is updated annually, information on the Fiscal Year 2017 criteria can be found here: https://www.ruralcenter.org/resource-library/flex-eligibility-criteria-for-mbqip-participation-and-waiver-template

• Geographic area and number and percentage of accountable entities and patients included: More than 1100 of the 1341 CAHS in the country currently report the EDTC measure through MBQIP. There are reporting CAHs in each of the 45 states that have CAHs (5 states do not have any CAHs). CAHs report at least 45 eligible cases each quarter. If they have more than 45 eligible cases they are allowed to use a sampling process, if they have less than 45 cases they report on all eligible cases.

• Level of measurement and setting: CAH Emergency Departments is the setting for the measure. CAHs collect data on individual patients and report numerators and denominators for the overall measure and each of the sub-measures on a quarterly basis to their state Flex program.

Minnesota Statewide Quality Reporting and Measurement System (SQRMS)

• Name of program and sponsor: Minnesota Statewide Quality Reporting and Measurement System (SQRMS), Minnesota Department of Health. Data is publically reported on the MN HealthScores website: http://www.mnhealthscores.org/

• Purpose: Minnesota's 2008 Health Reform Law requires the Commissioner of Health to establish a standardized set of quality measures for health care providers across the state. The goal is to create a uniform approach to quality measurement to enhance market transparency and drive health care quality improvement through an evolving measurement and reporting strategy. MN CAHs are not exempt from the SQRMS program (CAHs are exempt from Federal hospital reporting programs). The EDTC measure is a required SQRMS measure for CAHs in MN.

• Geographic area and number and percentage of accountable entities and patients included: All 78 MN CAHs reported the EDTC measure the most recent reporting quarter available (Q2 2017). CAHs report at least 45 eligible cases each quarter. If they have more than 45 eligible cases they are allowed to use a sampling process, if they have less than 45 cases they report on all eligible cases.

• Level of measurement and setting: CAH Emergency Departments is the setting for the measure. CAHs collect data on individual patients and report numerators and denominators for the overall measure and each of the sub-measures on a quarterly basis to their state Flex 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?*)
MBQIP has this measure included in its phase 3 reporting plan.

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

This measure is currently reported through MBQIP.

The Medicare Beneficiary Quality Improvement Project (MBQIP) is a quality improvement activity under the Medicare Rural Hospital Flexibility (Flex) grant program of the Health Resources and Services Administration's Federal Office of Rural Health Policy (FORHP). This is the process through which CAHs are required to report improvement measures.

https://www.ruralcenter.org/resource-library/flex-eligibility-criteria-for-mbqip-participation-and-waiver-templates

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.

Feedback of scores and comparison scores, national and local averages are provided as above.

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.

Reports are produced by a contractor to HRSA FORHP on a quarterly basis. Each quarter, the FORHP Project Officer for each state Flex program distributes the individual hospital and state summary data to the individual state Flex programs, who then in turn distribute the reports to the CAHs in their state. Individual state Flex programs provide support to CAHs on interpretation and use of the measure, but a guide on how to use the reports is also available: https://www.ruralcenter.org/resource-library/interpreting-mbqip-hospital-data-reports-for-quality-improvement. MN CAHs receive their data through this same process even though the measure is also used for public reporting through the MN SQRMS program.

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.

Reports are produced by a contractor to HRSA FORHP on a quarterly basis. Each quarter, the FORHP Project Officer for each state Flex program distributes the individual hospital and state summary data to the individual state Flex programs, who then in turn distribute the reports to the CAHs in their state. Individual state Flex programs provide support to CAHs on interpretation and use of the measure, but a guide on how to use the reports is also available: https://www.ruralcenter.org/resource-library/interpreting-mbqip-hospital-data-reports-for-quality-improvement. MN CAHs receive their data through this same process even though the measure is also used for public reporting through the MN SQRMS program.

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

Voluntary participants appreciate the feedback and have made simple process design improvements. Non-voluntary participants struggle with the indications of inadequate process completion.

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

Non-users have not been available to the measure sponsors. Some ideas may be that because CAHs have limited resources, their quality measurement resources may be focused on other issues. Some CAH may have improved their scores on these measures to near 100% and shifted their focus on other issues.

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.

Feedback and frequently asked questions are used for the annual update and clarification of specifications.

Tools developed by those being measured are sought by new reporters. Newsletters from Stratis Health contain clarifications and updates by the MBQIP monitor.

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.

Marked improvement is documented in all areas. Improvement is still needed for transfers to non-acute settings.

Improvement: Data and Trends (Q1 2015 – Q3 2016)

Every EDTC sub-measure and the composite EDTC-All measure has consistently improved between Q1 2015 and Q3 2016. Most markedly, EDTC-All has risen from 51.8% to 74.3% (a 22.5 percentage point increase).

Table 2: Timeframe

		EDTC-1 Percent	EDTC-2 Percent	EDTC-3 Percent	EDTC-4 Percent	EDTC-5 Percent	EDTC-6 Percent	EDTC-7 Percent	EDTC-All Percent
Q1	2015	84.5%	86.8%	87.6%	86.1%	84.1%	77.0%	90.2%	51.8%
Q2	2015	86.2%	89.6%	88.5%	87.2%	85.4%	78.9%	90.5%	56.8%
Q3	2015	90.4%	92.9%	92.1%	91.1%	89.8%	84.3%	94.8%	63.7%
Q4	2015	87.7%	90.2%	89.7%	88.4%	87.1%	81.9%	92.2%	65.6%
Q1	2016	92.8%	94.8%	93.8%	91.6%	91.5%	86.8%	94.7%	70.6%
Q2	2016	92.9%	93.6%	93.1%	91.0%	91.3%	86.6%	94.7%	72.2%
Q3	2016	93.2%	93.6%	93.5%	91.7%	91.6%	87.5%	94.9%	74.3%

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.

None. These measures occur after care is provided. It impacts communication of care provided and may improve care at the next facility.

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

None.

5. Comparison to Related or Competing Measures

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

5. Relation to Other NQF-endorsed Measures

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

No

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

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

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?
5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

5b. Competing Measures

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

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

Appendix

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

Attachment **Attachment:** Revised_EDTC_Measure_Specifications_Manual_Oct_2019-Final_-1-.pdf

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): University of Minnesota Rural Health Research Center

Co.2 Point of Contact: JILL, KLINGNER, KLIN0089@UMN.EDU, 218-726-8626-

Co.3 Measure Developer if different from Measure Steward: University of Minnesota Rural Health Research Center

Co.4 Point of Contact: JILL, KLINGNER, KLIN0089@UMN.EDU, 218-726-8626-

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.

All work group members reviewed the specifications and results related to the measure and results.

Ira Moscovice PhD Principal Investigator University of Minnesota RHRC

SHAILEY PRASAD MD FP Shailey Prasad MD FP

Tom Dean MD Rural FP Physician

John Supplitt AHA Rural Panel Director

Cathy Pfaff RN Rural Quality Improvement Consultant

Tim Size – Rural Health Network Director

Jennifer Lundblad - CEO Stratis Health QIO

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2003

Ad.3 Month and Year of most recent revision: 10, 2019

Ad.4 What is your frequency for review/update of this measure? ANNUALLY

Ad.5 When is the next scheduled review/update for this measure? 10, 2020

Ad.6 Copyright statement: X

Ad.7 Disclaimers: X

Ad.8 Additional Information/Comments: X