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 #: 0097

Corresponding Measures:

De.2. Measure Title: Medication Reconciliation Post-Discharge

Co.1.1. Measure Steward: National Committee for Quality Assurance

De.3. Brief Description of Measure: The percentage of discharges from January 1–December 1 of the measurement year for patients 18 years of age and older for whom medications were reconciled the date of discharge through 30 days after discharge (31 days total).

1b.1. Developer Rationale: The intent of this measure is to address a critical component of handoffs between inpatient and outpatient providers: reconciliation of medication lists. Incomplete or delayed communication between the inpatient facility and a patient's primary or ongoing care provider may result in duplication or omission of medications or the administration of medications with potentially harmful interactions. Timely reconciliation of the discharge medication list and the outpatient medical record medication list will reduce complications resulting from drug errors, interactions, omissions, or duplications for patients discharged from an inpatient facility.

S.4. Numerator Statement: Medication reconciliation conducted by a prescribing practitioner, clinical pharmacist or registered nurse, as documented through either administrative data or medical record review on the date of discharge through 30 days after discharge (31 total days).

S.6. Denominator Statement: All acute or nonacute inpatient discharges on or between January 1 and December 1 of the measurement year for patients who are 18 years and older.

S.8. Denominator Exclusions: No exclusions.

De.1. Measure Type: Process

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

S.20. Level of Analysis: Health Plan

IF Endorsement Maintenance – Original Endorsement Date: May 01, 2007 **Most Recent Endorsement Date:** Dec 10, 2015

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? N/A

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

The developer provided the following logic model:

- Patient is discharged from an acute or nonacute inpatient setting with a discharge medication list >>> Member is seen for post-discharge hospital follow-up>>> Medication reconciliation is conducted where a provider compares the discharge medication list and/or current medication list to the medication list in the outpatient record >>> Provider reviews medication lists with patient for appropriateness, potential side effects and adherence concerns >>>> Provider identifies if/when medications should be discontinued, changed, or added >>> Provider educates individual about medication list >>> Medication errors and adverse drug events are avoided

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

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

Updates:

- The developer conducted a 2020 review of the literature on the benefits of medication reconciliation, particularly for patients who are transferred between care facilities.
- The developer did not report any systematic reviews of the effect of medication reconciliation in the outpatient setting alone on health outcomes for adults. However, several studies have shown a decrease in medication errors when medication reconciliation, and other transition interventions, are implemented (Bayoumi 2009, Coleman 2003, Geurts 2012, Gillespie 2009, Midlov 2012, Nassaralla 2007).

- The developer noted the high prevalence of adverse drug events, and that about half of all adverse drug events are considered preventable (Agency for Healthcare Research and Quality 2019) and that on average 82% of adults in the U.S. take at least one medication and that 62% have multiple chronic conditions.
- The developer notes that poor hospital transitions are not only associated with poor health outcomes, but also increased health care utilization and cost, including duplicate medical services, medication errors and increased emergency department (ED) visits and readmissions (Sato 2011).
- The developer notes that hospital medication records are often incomplete when patients are admitted. A comparison of medication histories maintained for admitted patients with community pharmacy records revealed that hospital records omitted 25 percent of the medications in use. As a result, patients were discharged from the hospital without being continued on some chronic medications (Lau 2000).
- Significant changes can occur to a patient's medications during hospitalization. Beers et al. found that 45 percent of all discharge medications were initiated during hospitalization. Provider errors and patient misunderstanding of discharge medications are also common. One observational study found that 81.4 percent of patients experienced a provider error or had no understanding of at least one intended medication change upon discharge. Providers were more likely to make an error on a medication that was unrelated to the primary diagnosis, which emphasizes the importance of knowing the patient's current medications upon admission and discharge so that they are properly reconciled. Patients were more likely to misunderstand medication changes that were unrelated to the primary diagnosis, which stresses the importance of proper communication to the patient prior to and following discharge. (Ziaeian 2012)
- Resolving discrepancies in a patient's medication list reduces the risk of adverse drug interactions and helps physicians minimize duplication and complexity of a medication regimen (Wenger 2004).

Questions for the Committee:

- The evidence provided by the developer is somewhat updated, but largely the same as last cycle. Does the Committee agree there is no need for repeat discussion and vote on Evidence?
- For structure, process, and intermediate outcome measures:
    - 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?
- For possible exception to the evidence criterion:
    - Are there, or could there be, performance measures of a related health outcome, OR evidence-based intermediate clinical outcomes, intervention/treatment?
    - Is there evidence of a systematic assessment of expert opinion beyond those involved in developing the measure?
    - Does the SC agree that it is acceptable (or beneficial) to hold providers accountable without empirical evidence?

Guidance from the Evidence Algorithm

Box 1 -> Box 3 -> Box 7 -> Box 10 -> Yes - INSUFFICIENT

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

RATIONALE: A performance measure outcome could be actual medication errors. There is no systematic review of the literature on this topic. There is no clear evidence that links medication reconciliation to improved outcomes using rigorously designed studies.

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.

The following data were extracted from HEDIS data collection for Medicare Health Plans. Performance data is summarized at the health plan level and summarized by mean, standard deviation, and performance at the 10th, 25th, 50th, 75th and 90th percentile.

Table 1. Variation in Performance Across Medicare Health Plans

YEAR	N	MEAN RATE	SD	MIN	10TH	25TH	50TH	75TH	90TH	MAX
2018*	472	61.3%	18.0%	0.0%	36.0%	50.5%	63.0%	74.6%	83.7%	97.8%
2017	465	52.7%	20.0%	0.0%	25.0%	38.7%	54.0%	68.9%	77.6%	97.1%
2016^	467	46.4%	20.6%	0.0%	16.8%	31.4%	47.0%	61.3%	73.4%	98.0%

*For 2018 the average denominator was 504, with a standard deviation of 2,425.

Disparities

No disparities data are reported. The developer describes a rationale why disparities data should not be included.

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, 1b, 1c)

1a. Evidence to Support Measure Focus: 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.

- This important aspect of clinical care was emphasized with an updated literature review. I'm new to the committee but was struck by inclusion in the numerator for patients up to 31 days following discharge.
- Insufficient evidence with exception. The supporting literature is old. Consider Daliri, S., Bouhnouf, M., van de Meerendonk, H. W., Buurman, B. M., op Reimer, W. J. S., Kooij, M. J., & Karapinar–Çarkit, F. (2020). Longitudinal medication reconciliation at hospital admission, discharge and post-discharge. Research in Social and Administrative Pharmacy. The greater problem may be the inpatient stay and discharge. Singular focus on post-discharge may be less impactful.
- Clear review of evidence on risks of medication interactions after hospitalization. Medication reconciliation alone has not been tested but included in successful care transition interventions cited.
- Worthy of a discussion as noted in prelim review
- Evidence exists that medication reconciliation reduces medication discrepancies and medication errors that result in adverse events when performed appropriately and with competence.
- Limited evidence and measure will not help with medication errors.
- The evidence is Insufficient – i.e. no systematic review for a process measure. Although intuitive and there's some evidence (not systematic review) that medication reconciliation reduces medication errors, often the study included other transition interventions and "medication reconciliation" studied in the

literature was not operationalized as what is proposed here in the measure. For example, one study the intervention included clinic calling patient ahead of appointment to remind them to bring updated list of medications then having patient filling out a form about medications etc. I can imagine that if patient did not bring meds and cannot recall meds then there's still room for error even if the clinician tried to reconcile medications after discharge. However this is a maintenance measure if I understand correctly, so I would like to understand better how it passed this criterion in the past?

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?

- Yes, and there is a gap in care
- Support high rating. However disparity may lie in the lack of standardized approaches, processes and tools in any stage of the continuum of care.
- Significant variability noted on outcomes from 2018 reported, indicating gaps in care.
- High gap
- This measure reports whether medication reconciliation was performed and the timeliness of performance and a post-discharge performance gap is noted.
- Performance shows improvement 2016-2018.
- As of 2018, out of 472 health plans, variation range from 36% at 10th percentiles to 83.7% at 90th percentile, showing significant variation and opportunity for improvement. There was no data on subgroups or demonstrated disparities. Rating is High.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

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

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

Reliability

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

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

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

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

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

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

Reliability

- Reliability testing was performed using 2018 HEDIS data. Results are below.

Table 2a. Mean Signal-To-Noise Reliability, Standard Error (SE) and 95% Confidence Interval (95% CI) for the *Medication Reconciliation Post-Discharge* Measure by Terciles of the Denominator Size and for All Submissions, 2018

Stratification	Number of Plans	Number of Eligible Members per Plan (min max)	Mean Signal To Noise Reliability	SE	95% CI
All Medicare	**472**	**32 - 51839**	**0.977**	**0.001**	**(0.975, 0.979)**
Tercile 1	128	32 - 410	0.964	0.003	(0.959, 0.970)
Tercile 2	339	411 - 411	0.981	0.000	(0.981, 0.982)
Tercile 3	5	1442 - 51839	1.000	0.000	(0.999, 1.000)

SE: Standard Error of the mean.

95% CI: 95% confidence interval.

Table 2b. Distribution of Plan-Level Signal-To-Noise Reliability for the *Medication Reconciliation Post-Discharge* Measure by Terciles of the Denominator Size and for All Submissions, 2018

Stratification	Number of Plans	Min	Distribution of Plan Estimates of Signal to Noise Reliability: P10	Distribution of Plan Estimates of Signal to Noise Reliability: P25	Distribution of Plan Estimates of Signal to Noise Reliability: P50	Distribution of Plan Estimates of Signal to Noise Reliability: P75	Distribution of Plan Estimates of Signal to Noise Reliability: P90	Distribution of Plan Estimates of Signal to Noise Reliability: Max
All Medicare	**472**	**0.804**	**0.961**	**0.981**	**0.982**	**0.986**	**0.990**	**1.000**
Tercile 1	128	0.832	0.922	0.955	0.976	0.984	0.988	1.000
Tercile 2	339	0.977	0.977	0.978	0.979	0.983	0.988	0.997
Tercile 3	5	0.999	0.999	1.000	1.000	1.000	1.000	1.000

Validity

- The developers tested for construct (empirical) validity comparing medication reconciliation post-discharge to three other HEDIS measures.

Measure	**Correlation Coefficient:** Notification of Inpatient Admission	**Correlation Coefficient:** Receipt of Discharge Information	**Correlation Coefficient:** Patient Engagement After Inpatient Discharge Rate
Medication Reconciliation Post-Discharge	0.45	0.43	0.60
(N=, p value =)	(449, $p < 0.001$)	(449, $p < 0.001$)	(472, $p < 0.001$)

- Face validity results
    - The developer reported that their MAPs agreed with the measure's intent and proposed specification, the majority of public comments received supported the measure, and our CPM, and subsequently our Board of Directors, approved the measure for HEDIS reporting.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?

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

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

Committee Pre-evaluation Comments:

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability-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?

- No concerns
- Moderate to high. I don't understand how the authors are measuring the link between number of patients/plans and inappropriate use of prescriptions. Limited focus on actual outcomes of medication reconciliation.
- Reliability testing was strong. Estimates of signal-noise reliability were high across tertiles.
- High reliability
- No comments
- none
- Reliability data demonstrates good reliability with reliability estimate of 0.977. No concerns. Rating High.

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

- None
- No, not the existing measure.
- No
- no
- No comments
- no
- No concerns.

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

- None
- Moderate validity.
- No. Good correlations with related measures suggests this one is assessing a robust construct with face validity and construct validity
- no
- No comments
- Measure does not go far enough in ensuring that quality of care is being assured. Reconciliation should include an expert assessment of whether medications are optimized, more than replicated between care facilities.
- I'm not sure that correlation with other transitional care measures necessarily support the construct validity of this measure. I think face validity is present in that it's trying to measure medication reconciliation, back to the evidence question it's not clear if measuring medication reconciliation will necessarily improve patient safety or outcomes. Rating is Moderate.

2b2-3. 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 concerns
- N/A
- No exclusions (would patients with 0 prescribed medications need this visit?)..how are post-acute stays accounted for if patient stays >30 days in a nursing home or rehab hospital after discharge?
- No concerns with validity
- Not applicable
- ok
- No risk adjustment which is appropriate. Exclusion includes those patients on hospice, which is appropriate.

2b4-7. 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?

- Concern over the length of time med rec can take and still be included
- Prior testing data is presented differently.
- Clinically meaningful differences in number of patients/plan who did not have reconciliation of medications after hospitalization
- No concerns with threats
- This measure reports whether med rec was done or not but does not imply quality.
- Ideally, the measure 'Patient Engagement After Inpatient Discharge' could be combined with this measure because verification of patient understanding of her medications is an essential part of reconciliation.

- 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.
    - Some data elements are in defined fields in electronic sources. Health plans and providers that use an electronic health record to capture medication reconciliation use that data to report on this 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?**

- No concerns from 2015 update
- Moderate feasibility- Workflow issue, data collection and integrity is vulnerable to human error and variances in the patient's ability to participate in the process.
- Feasible if leveraging EMR data with common elements. If reconciliation is documented in free text it may be harder to accurately capture.
- Moderate feasibility
- Measure is feasible
- none
- Already being implemented. It's possible that medication reconciliation is done but not well documented, but I think that's less common.

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**

Accountability program details

Public Reporting

Annual State of Health Care Quality

https://www.ncqa.org/report-cards/health-plans/state-of-health-care-quality-report/

Health Plan Ratings

http://www.ncqa.org/ReportCards/HealthPlans/HealthInsurancePlanRankings/HealthPlanRatingsPreview.aspx

Annual State of Health Care Quality

https://www.ncqa.org/report-cards/health-plans/state-of-health-care-quality-report/

Health Plan Ratings

http://www.ncqa.org/ReportCards/HealthPlans/HealthInsurancePlanRankings/HealthPlanRatingsPreview.aspx

Payment Program

Physician Quality Reporting Systems (PQRS)

http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/

Regulatory and Accreditation Programs

HEDIS®-Health Plan

http://www.ncqa.org/Programs/Accreditation/HealthPlanHP.aspx

HEDIS®-Health Plan

http://www.ncqa.org/Programs/Accreditation/HealthPlanHP.aspx

Quality Improvement (external benchmarking to organizations)

Quality Compass

https://www.ncqa.org/programs/data-and-information-technology/data-purchase-and-licensing/quality-compass/

CMS Medicare Advantage Plan Rating System (STARS)

https://www.medicare.gov/find-a-plan/questions/home.aspx

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

Health plans that report HEDIS calculate their rates and know their performance when submitting to NCQA. NCQA publicly reports rates across all plans and also creates benchmarks in order to help plans understand how they perform relative to other plans.

Additional Feedback:

NCQA states that measures are evaluated regularly. During this "reevaluation" process, NCQA seek broad input on the measure, including input on performance and implementation experience. NCQA uses several methods to obtain input, including vetting of the measure with several multi-stakeholder advisory panels, public comment posting, and review of questions submitted to the Policy Clarification Support System via MyNCQA.org. This information enables NCQA to comprehensively assess a measure's adherence to the HEDIS Desirable Attributes of Relevance, Scientific Soundness and Feasibility.

***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 The 2016 to 2018 data shows that while performance rates for this measure are low, they have increased in the past year (see section 1b.2 for summary of data from health plans). In 2018, the average performance was 61.3%. There was a 47.7 percentage point difference between plans at the 10th and 90th percentiles. This large difference in performance represents a persistent gap in care and room for improvement in medication reconciliation for health plan members who use prescription medications. As noted in a prior section, 2016 average performance was 46.4 and 2018 average performance was 61.3, possibly reflecting revisions in the measure specification. The number of plans reporting has slightly increased over the years from 2016 (n=467) to 2018 (n=472). Overall performance has also increased over the years.

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 None reported

Potential harms None reported

***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

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?

- Used for publicly reported ratings. Very important for clinical care and reviewed internally by systems as well as health plans.
- Pass
- Public reporting of data is ongoing, and currently being used in an accountability program.
- In use
- This is an "entry level" type of medication reconciliation process measure. Better that time is utilized on ensuring post discharge medication reconciliation is performed AND is of high quality.
- ok
- Currently being used and publicly reported on reports rating health plans. The health plans are informed on their ratings and have been given opportunity to provide feedback. Rating Pass.

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.

- No additional concerns
- The benefits of the measure outweigh risks.
- Benefits outweigh harms, and no unintended consequences seem to have occurred with implementation
- Benefits outweigh any harms
- This measure can only be useful if applied concomitantly with a measure of actual med rec quality, the only one currently approved that measures quality is #2456 medication discrepancies
- The real question is whether medications have been prescribed based on indications, not whether they are consistent from facility to facility. The latter does provide some protection against harm to patient.
- Has led to improvement in the rates over the years from 46.4% in 2016 to 61.3%. No identified unintended consequences. Rating high.

Criterion 5: Related and Competing Measures

Related or competing measures

0419 : Documentation of Current Medications in the Medical Record
0553 : Care for Older Adults (COA) – Medication Review
2456 : Medication Reconciliation: Number of Unintentional Medication Discrepancies per Medication Per Patient
2988 : Medication Reconciliation for Patients Receiving Care at Dialysis Facilities
3317 : Medication Reconciliation on Admission

Harmonization

Developer reports that this measure is harmonized.

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

5. 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?

- Appears to be harmonized with other measures
- The measure is justified. Do not see a measure for med-rec on discharge.
- A number are related or competing, but this has been harmonized with others
- Several, noted in prelim document
- Recommend that this measure be bundled with #2456 medication discrepancies to make them useful to providers
- yes
- I find the explanations provided to differentiate this from other related measures to be satisfactory.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 01/15/2021

- No NQF Members have submitted support/non-support choices as of this date.
- No Public or NQF Member comments submitted as of this date.

Scientific Acceptability: Preliminary Analysis Form

Measure Number: Insert measure number here

Measure Title: Insert measure title here

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_xxxx" document, items S.1-S.22

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

2. **Briefly summarize any concerns about the measure specifications.**

   None

RELIABILITY: TESTING

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

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Signal-to-noise testing was done. Data were reported with 95% CI for all plans and stratified by plan size. The developer also reported the distribution of scores. These methods are appropriate.

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

   Submission document: The reliability estimate was 0.977, 95% CI is (0.975, 0.979) which indicates very good reliability. Stratified analyses show that reliability increases as plan size gets larger. The distribution of reliability scores was also narrow, which further indicates good reliability for this measure.

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 reliability testing used the correct methods and the results demonstrated excellent reliability (near 1) with tight confidence intervals.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: No concerns.

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

    Submission document: No concerns.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: There is only a single data source.

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

    Submission document: There are no missing data.

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:** No risk adjustment.

For cost/resource use measures ONLY:

17. **Are the specifications in alignment with the stated measure intent?**

    ☐ **Yes** ☐ **Somewhat** ☐ **No (If "Somewhat" or "No", please explain)**

18. **Describe any concerns of threats to validity related to attribution, the costing approach, carve outs, or truncation (approach to outliers):**

VALIDITY: TESTING

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

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

☒ **Face validity**

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

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

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

Submission document: Methods are appropriate for construct validity. Face validity approach is reasonable.

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

Submission document: Test sample is adequate. Results demonstrate validity. Does demonstrate association between three other similar measures with positive, statistically significant correlations.

23. **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)

24. **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)

25. **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.)

26. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.** Construct validity testing demonstrate an association with similar, related metrics with $p<0.001$. Face validity results are supportive of this measure.

Developer Submission

NQF #: 0097

Corresponding Measures:

De.2. Measure Title: Medication Reconciliation Post-Discharge

Co.1.1. Measure Steward: National Committee for Quality Assurance

De.3. Brief Description of Measure: The percentage of discharges from January 1–December 1 of the measurement year for patients 18 years of age and older for whom medications were reconciled the date of discharge through 30 days after discharge (31 days total).

1b.1. Developer Rationale: The intent of this measure is to address a critical component of handoffs between inpatient and outpatient providers: reconciliation of medication lists. Incomplete or delayed communication between the inpatient facility and a patient's primary or ongoing care provider may result in duplication or omission of medications or the administration of medications with potentially harmful interactions. Timely reconciliation of the discharge medication list and the outpatient medical record medication list will reduce complications resulting from drug errors, interactions, omissions, or duplications for patients discharged from an inpatient facility.

S.4. Numerator Statement: Medication reconciliation conducted by a prescribing practitioner, clinical pharmacist or registered nurse, as documented through either administrative data or medical record review on the date of discharge through 30 days after discharge (31 total days).

S.6. Denominator Statement: All acute or nonacute inpatient discharges on or between January 1 and December 1 of the measurement year for patients who are 18 years and older.

S.8. Denominator Exclusions: No exclusions.

De.1. Measure Type: Process

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

S.20. Level of Analysis: Health Plan

IF Endorsement Maintenance – Original Endorsement Date: May 01, 2007 **Most Recent Endorsement Date:** Dec 10, 2015

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? N/A

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

MRP_nqf_evidence_attachment_7.1-637396684764472863.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)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0097

Measure Title: Medication Reconciliation Post-Discharge

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

Date of Submission: **11/2/2020**

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

1a. Evidence to Support the Measure Focus

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

- Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.
4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.
5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.
6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

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

Outcome

☐ Outcome:

☐Patient-reported outcome (PRO):

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

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

☒ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

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

2020 Submission

Patient is discharged from an acute or nonacute inpatient setting with a discharge medication list >>> Member is seen for post-discharge hospital follow-up>>> Medication reconciliation is conducted where a provider compares the discharge medication list and/or current medication list to the medication list in the outpatient record >>> Provider reviews medication lists with patient for appropriateness, potential side effects and adherence concerns >>>> Provider identifies if/when medications should be discontinued, changed, or added >>> Provider educates individual about medication list >>> Medication errors and adverse drug events are avoided.

2015 Submission

Patient is discharged from an acute or nonacute inpatient setting with a discharge medication list >>> Medication reconciliation is conducted where a provider compares the discharge medication list to the medication list in the outpatient record >>> Provider reviews medication list with patient for appropriateness, potential side effects and adherence concerns >>>> Provider identifies where medications should be discontinued, changed, or added >>> Provider educates individual about medication list >>> Medication errors and adverse drug events are avoided.

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.)

2020 Submission

N/A

2015 Submission

Did not answer

****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.

2020 Submission

Not an outcome measure

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

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

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

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

☐ Other

Systematic Review

Evidence

Source of Systematic Review:

- Title
- Author
- Date
- Citation, including page number
- URL

*

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.

*

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade

*

Provide all other grades and definitions from the evidence grading system

*

Grade assigned to the **recommendation** with definition of the grade

*

Provide all other grades and definitions from the recommendation grading system

*

Body of evidence:

- Quantity – how many studies?
- Quality – what type of studies?

*

Estimates of benefit and consistency across studies

*

What harms were identified?

*

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?

*

*cell intentionally left blank

_____

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. A list of references without a summary is not acceptable.

2020 Submission

Studies consistently point towards the benefits of performing medication reconciliation, particularly for patients who are transferred between care facilities, which increases the risk for medication discrepancies in the patient's medication regimen. Medication reconciliation post-discharge can catch potentially harmful omissions or changes in prescribed medications, particularly for older adults and disabled patients, who are prescribed a greater quantity and variety of medications (Leape 1991). To our knowledge, there are no systematic reviews of the effect of medication reconciliation in the outpatient setting alone on health outcomes for adults. However, individual studies have shown a decrease in medication errors when medication reconciliation, and other transition interventions, are implemented (Bayoumi 2009, Coleman 2003, Geurts 2012, Gillespie 2009, Midlov 2012, Nassaralla 2007).

The high prevalence of prescription medications can result in potentially negative consequences for patients if not used and monitored appropriately. Approximately 1.5 million preventable adverse drug events occur in the United States each year (Johnson 1995). Many of these result from medication errors, drug interactions or inappropriate use of medications. About half of all adverse drug events are considered preventable (Agency for Healthcare Research and Quality 2019).

On average, 82 percent of adults in the U.S. take at least 1 medication (prescription or nonprescription, vitamin/mineral, herbal/natural supplement); 29 percent take 5 or more. Older adults are the biggest consumers of medications: 17 percent–19 percent of people 65 and older take at least 10 medications in a given week (Slone Survey 2014).

62 percent of adults 65 and older have multiple chronic conditions; the higher number of chronic conditions they experience, the more providers are involved in their care. As the number of providers increases, the less likely patients are to understand, remember and reconcile multiple instructions (Vogeli 2007).

Financial Impact:

Poor hospital transitions are not only associated with poor health outcomes, but also increased health care utilization and cost, including duplicate medical services, medication errors and increased emergency department (ED) visits and readmissions (Sato 2011).

With hospital stays costing the U.S. $377.5 billion per year, and Medicare members contributing to increased lengths of stay, there is more pressure for hospitals to improve their delivery of care and lower patient harm. Part of this includes examining the discharge process, particularly for Medicare members, to prevent further variability in discharge practices that could result in re-hospitalization and ED visits (Health Catalyst 2017).

Evidence:
Hospital medication records are often incomplete when patients are admitted. A comparison of medication histories maintained for admitted patients with community pharmacy records revealed that hospital records omitted 25 percent of the medications in use. As a result, patients were discharged from the hospital without being continued on some chronic medications (Lau 2000).

Significant changes can occur to a patient's medications during hospitalization. Beers et al. found that 45 percent of all discharge medications were initiated during hospitalization. Provider errors and patient misunderstanding of discharge medications are also common. One observational study found that 81.4 percent of patients experienced a provider error or had no understanding of at least one intended medication change upon discharge. Providers were more likely to make an error on a medication that was unrelated to the primary diagnosis, which emphasizes the importance of knowing the patient's current medications upon admission and discharge so that they are properly reconciled. Patients were more likely to misunderstand medication changes that were unrelated to the primary diagnosis, which stresses the importance of proper communication to the patient prior to and following discharge. (Ziaeian 2012)

Resolving discrepancies in a patient's medication list reduces the risk of adverse drug interactions and helps physicians minimize duplication and complexity of a medication regimen (Wenger 2004).

2015 Submission

Studies consistently point towards the benefits of performing medication reconciliation, particularly for patients who are transferred between care facilities, which increases the risk for medication discrepancies in the patient's medication regimen. To our knowledge, there are no systematic reviews of the effect of medication reconciliation alone on health outcomes for adults. However, individual studies have demonstrated a decrease in medication errors when medication reconciliation among other care transition interventions are implemented (Bayoumi 2009; Coleman 2003; Gillespie 2009; Nassaralla 2007; Geurts 2012; Midlov 2012). Medication reconciliation is a critical component of several widely disseminated care transitions models including the Transitional Care Model, Care Transitions Program, Project RED, and Project BOOST.

Medication reconciliation post-discharge is an important step to catch potentially harmful omissions or changes in prescribed medications, particularly in elderly patients that are prescribed a greater quantity and variety of medications (Leape 1991). Hospital admissions are associated with unintentional discontinuation of medication for chronic conditions (Bell 2011) and medication errors (Stafford 2011; IOM 2006). Although the magnitude of the effect of medication reconciliation alone on patient outcomes is not well studied, there is agreement among experts that potential benefits outweigh the harm (Coleman 2003; Pronovost 2003;IOM 2002; IOM 2006). Medication reconciliation post-discharge is recommended by the Joint Commission patient safety goals (Kienle 2008), the American Geriatric Society (Coleman 2003), Society of Hospital Medicine (Kripalani 2007; Grennwald 2010), ACOVE (Assessing Care of Vulnerable Elders; Knight 2001),and the Task Force on Medicines Partnership (2005). Additionally, measurement of medication reconciliation post-discharge has been cited by the National Quality Forum and the National Priorities Partnership as a measurement priority area (NQF 2010).

Quality of Evidence:

Medication reconciliation post-discharge is widely regarded as good practice. Interventions that have targeted reducing adverse medication events have combined medication reconciliation with other care coordination and transition interventions. Therefore, the body of evidence directly linking medication reconciliation with patient outcomes is moderate. While all studies have shown a positive effect of medication reconciliation on reducing medication errors, very few have had the power to show an effect on outcomes such as morbidity and mortality. Despite this limitation, there is general expert consensus that the benefits of medication reconciliation outweigh the harms.

Estimate of Benefit and Consistency of Results across Studies:

All studies have shown a positive effect of medication reconciliation on reducing medication errors. Studies have shown mixed results when examining the effect of medication reconciliation on morbidity and mortality. No studies have shown any harm to the patient from medication reconciliation. In one study, the percentage of patients affected by adverse drug events fell from 36.9% to 9.3% with the use of medication reviews (IOM 2011). This intervention may also ease the financial burden that medication errors place on the medical system. A study utilizing a pharmacist-led medication review concluded that there was a 16% reduction in all visits to the hospital, a 47% reduction in visits to the emergency department, and $230 decrease in cost per patient (Gillespie 2009).

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

2020 Submission

A search of relevant literature was conducted to identify studies on outcomes of medication reconciliation, the financial impact of poor hospital transitions when medication reconciliation is important and preventable burden on the healthcare system when poor transitions and/or medication reconciliation occur.

2015 Submission

A search of relevant literature was conducted to identify studies on outcomes of medication reconciliation.

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

2020 Submission

Agency for Healthcare Research and Quality (AHRQ). 2019. *Medication Errors and Adverse Drug Events.* https://psnet.ahrq.gov/primers/primer/23/Medication-Errors-and-Adverse-Drug-Events (Accessed August 27, 2019).

Bayoumi, I., M. Howard, A.M. Holbrook, I. Schabort. 2009. "Interventions to improve medication reconciliation in primary care." *Ann Parmacother* 43:1667-75.

Beers, M.H., J. Dang, J. Hasegawa, I.Y. Tamai. 1989. "Influence of hospitalization on drug therapy in the elderly." J Am Geriatr Soc 37(8):679-83.

Coleman, E.A., C.E. Boult, American Geriatrics Society Health Care Systems Committee. 2003. "Improving the Quality of Transitional Care for Persons with Complex Care Needs." *J Am Geriatr Soc*. 51(4):556-7.

Geurts, M.M., J. Talsma, J.R. Brouwers, J.J. de Gier. 2012. "Medication Review and Reconciliation with Cooperation between Pharmacist and General Practitioner and the Benefit for the Patient: a Systematic Review." *Br J Clin Pharmacol*. Epub ahead of print. Jan 13.

Gillespie U, Alassaad A, Henrohn D, et al. A Comprehensive Pharmacist Intervention to Reduce Morbidity in Patients 80 Years or Older. *Arch Intern Med.* 2009;169:894-900.

Health Catalyst. 2017. *Patient-Centered LOS Reduction Initiative Improves Outcomes, Lowers Costs.* https://downloads.healthcatalyst.com/wp-content/uploads/2016/06/Patient-Centered-LOS-Reduction-Initiative-Improves-Outcomes-Lowers-Costs.pdf (Accessed August 27, 2019).

Johnson, J.A., and J.L. Bootman. 1995. "Drug-related morbidity and mortality: A cost-of-illness model." *Arch Intern Med* 155:1949–56.

Lau, H.S., C. Florax, A.J. Porsius, A. De Boer. 2000. "The completeness of medication histories in hospital medical records of patients admitted to general internal medicine wards." *Br J Clin Pharmacol* 49(6):597-603.

Leape, L.L., T.A. Brennan, et al. 1991. "The Nature of Adverse Events in Hospitalized Patients." *N Engl J Med* 324(6):377-84.

Midlov, P., L. Bahrani, M. Seyfali et al. 2012. "The effect of medication reconciliation in elderly patients at hospital discharge." *Int J Clin Pharm* 34(1):113-9. doi: 10.1007/s11096-011-9599-6. Epub 2011 Dec 30.

Nassaralla, C.L., J.M. Naessens, R. Chaudhry, et al. 2007. "Implementation of a medication reconciliation process in an ambulatory internal medicine clinic." *Qual Saf Health Care* 16(2):90-4.

Patterns of medications use in the United States 2006: a report from the Slone Survey. http://www.bu.edu/slone/files/2012/11/SloneSurveyReport2006.pdf (Accessed July 17, 2014)

Sato, M., T. Shaffer, A.I. Arbaje and I.H. Zuckerman. 2011. "Residential and health care transition patterns among older Medicare beneficiaries over time." *The Gerontologist* 51(2), 170–8.

Vogeli, C., A.E. Shields, T.A. Lee, et al. 2007. "Multiple Chronic Conditions: Prevalence, Health Consequences, and Implications for Quality, Care Management, and Costs." *J Gen Intern Med* 22(suppl 3): 391–5.

Wenger, N.S. and R. Young. 2004. "Working paper: Quality Indicators of Continuity and Coordination of Care for Vulnerable Elder Persons." Rand.

Ziaeian, B., K.L.B. Araujo, P. Van Ness, L. Horwitz. 2012. "Medication Reconciliation Accuracy and Patient Understanding of Intended Medication Changes on Hospital Discharge." *J Gen Intern Med* 27(11):1513-20. doi: 10.1007/s11606-012-2168-4. Epub 2012 Jul 14.

2015 Submission

Citations:

Bayoumi I, Howard M, Holbrook Am, Schabort I. Interventions to improve medication reconciliation in primary care. Ann Parmacother. 2009; 43:1667-75.

Bell CM, Brener SS, Gunraj N, Huo C, Bierman AS, Scales DC, Bajcar J, Szarenstein M, Urbach DR. Association of ICU or hospital admission with unintentional discontinuation of medication for chronic diseases. JAMA. 2011;306:840-7.

Coleman EA, Boult CE on behalf of the American Geriatrics Society Health Care Systems Committee. Improving the Quality of Transitional Care for Persons with Complex Care Needs. Journal of the American Geriatrics Society. 2003;51(4):556-557.

Geurts MM, Talsma J, Brouwers JR, de Gier JJ. Medication Review and Reconciliation with Cooperation between Pharmacist and General Practitioner and the Benefit for the Patient: a Systematic Review. Br J Clin Pharmacol. 2012. Epub ahead of print. Jan 13.

Gillespie U, Alassaad A, Henrohn D, et al. A Comprehensive Pharmacist Intervention to Reduce Morbidity in Patients 80 Years or Older. Arch Intern Med. 2009;169:894-900.

Grennwald JL, Halasyamani L, Greene J, et al. Making inpatient medication reconciliation patient centered, clinically relevant, and implementable: A consensus statement on key principles and necessary first steps. J Hosp Med. 2010;5:477-85.

Institute of Medicine. Committee on Quality Health Care in America. Washington, DC: National Academy Press. 2002.

Institute of Medicine. Preventing Medication Errors. National Academies Press, Washington D.C. 2006.

Kienle P, Uselton JP. Maintaining Compliance with Joint Commission Medication Management Standards. Patient Safety and Quality Healthcare. 2008; July/August.

Kripalani S, Jackson AT, Schnipper JL, Coleman EA. Promoting effective transitions of care at hospital discharge: a review of key issues for hospitals. J Hosp Med. 2007;2:314-23.

Knight EL, Avorn J. Quality Indicators for appropriate medication use in vulnerable elders. Ann Intern Med. 2001:703-10.

Leape LL, Brennan TA, et al. The Nature of Adverse Events in Hospitalized Patients. N Engl J Med. 1991;324(6):377-84.

Midlov P, Bahrani L, Seyfali M, Hoglund P, Rickhag E, Eriksson T. The effect of medication reconciliation in elderly patients at hospital discharge. Int J Clin Pharm. 2012. Epub ahead of print. Feb 2012.

Mueller SK, Cunningham K, Kripalani S, Schnipper J. 2012. "Hospital-Based Medication Reconciliation Practices." Arch Intern Med. 2012;172(14):1057-1069 doi:10.1001/archinernmed.2012.2246.

Nassaralla CL, Naessens JM, Chaudhry R, et al. Implementation of a medication reconciliation process in an ambulatory internal medicine clinic. Qual Saf Health Care 2007; 16: 90-94.

National Quality Forum (NQF), Preferred Practices and Performance Measures for Measuring and Reporting Care Coordination: A Consensus Report, Washington, DC:NQF; 2010.

Pronovost P, Weast B, Schwarz M, et al. Medication Reconciliation: A Practical Tool to Reduce the Risk of Medication Errors. J Crit Care. 2003 Dec;18(4):201-5.

Stafford L, Stafford A, Hughes J, Angley M, Berezniki L, Peterson G. Drug-related problems identified in post-discharge medication reviews for patients taking warfarin. Int J Clin Pharm. 2011;33:621-6.

Task force on Medicines Partnership. The National Collaborative medicines Management Services Programme. Room for Review. A Guide to Medication Review. London, 2002. Accessed via: **http://www.medicines-partnership.org/medication-review**, Reviewed September 2005.

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.

The intent of this measure is to address a critical component of handoffs between inpatient and outpatient providers: reconciliation of medication lists. Incomplete or delayed communication between the inpatient facility and a patient's primary or ongoing care provider may result in duplication or omission of medications or the administration of medications with potentially harmful interactions. Timely reconciliation of the discharge medication list and the outpatient medical record medication list will reduce complications resulting from drug errors, interactions, omissions, or duplications for patients discharged from an inpatient facility.

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 following data are extracted from HEDIS data collection for Medicare Health Plans. Performance data is summarized at the health plan level and summarized by mean, standard deviation, and performance at the 10th, 25th, 50th, 75th and 90th percentile.

Table 1. Variation in Performance Across Medicare Health Plans

YEAR | N | MEAN RATE | SD | MIN | 10TH | 25TH | 50TH | 75TH | 90TH | MAX

2018* | 472 | 61.3% | 18.0% | 0.0% | 36.0% | 50.5% | 63.0% | 74.6% | 83.7% | 97.8%

2017 | 465 | 52.7% | 20.0% | 0.0% | 25.0% | 38.7% | 54.0% | 68.9% | 77.6% | 97.1%

2016^ | 467 | 46.4% | 20.6% | 0.0% | 16.8% | 31.4% | 47.0% | 61.3% | 73.4% | 98.0%

*For 2018 the average denominator was 504, with a standard deviation of 2,425.

^ Note: There is a change in trending for these results as they are based on a previous specification of the HEDIS measure. For 2016 this measure was updated to add Medicare as a product line and expand the age range to include Medicare patients 18 years and older.

The data referenced are extracted from HEDIS data collection reflecting the most recent years of measurement for this measure. In 2018, HEDIS measures covered more than 21 million Medicare enrollees. Below is a description of the denominator for this measure. It includes the number of health plans included in HEDIS data collection and the median and mean denominator for the measure across health plans.

YEAR | N Plans | Median Denominator Size per plan | Mean Denominator Size per plan

2016 | 467 | 411 | 582

2017 | 465 | 411 | 526

2018 | 472 | 411 | 504

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.

N/A

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.*

HEDIS data are stratified by type of insurance (e.g., Commercial, Medicaid, Medicare). While not specified in the measure, this measure can also be stratified by demographic variables, such as race/ethnicity or socioeconomic status, in order to assess the presence of health care disparities, if the data are available to a

plan. NCQA is actively engaged with partners including the CMS Office of Minority Health in identifying feasible methods to further integrate social risk factors into health plan quality measures, with a focus on stratification. Our work is aligned with recent recommendations from MedPAC and ASPE on optimal methods for addressing social risk in quality measurement and programs.1,2 This is an NCQA wide initiative. Our intent is to implement methods to bridge data concerns in the future.

HEDIS includes two measures that can be used as tools for assessing race/ethnicity and language needs of a plan's population: Race/Ethnicity Diversity of Membership and the Language Diversity of Membership. These measures promote standardized methods for collecting these data and follow Office of Management and Budget and National Academy of Medicine guidance for collecting and categorizing race/ethnicity and language data. In addition, NCQA's Multicultural Health Care Distinction Program outlines standards for collecting, storing, and using race/ethnicity and language data to assess health care disparities.

1. Medicare Payment Advisory Commission. (2020). The Medicare Advantage program: Status report. In Report to the Congress: Medicare Payment Policy (p. 397). http://medpac.gov/docs/default-source/reports/mar20_medpac_ch13_sec.pdf
2. Office of the Assistant Secretary for Planning and Evaluation, & U.S. Department of Health & Human Services. (2020). Second Report to Congress on Social Risk and Medicare's Value-Based Purchasing Programs. https://aspe.hhs.gov/social-risk-factors-and-medicares-value-basedpurchasing-programs

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

While there is little evidence on disparities in receiving medication reconciliation, we do see in the April 2019 Racial, Ethnic, and Gender Disparities in Health Care in Medicare Advantage Report from CMS that differences persist. For example, Blacks were less likely to have their medication reconciled within 30 days after being discharged from an inpatient facility (54.5% versus 58.0% for Whites). For Hispanics, they are more likely to have their medication reconciled than Whites when discharged from an inpatient facility (60.4% versus 58.0%). This data demonstrates that medication reconciliation varies based on race and ethnic group.

We do see sufficient evidence that disparities in health outcomes may be explained by a patients' inability to afford prescription medications. Studies have documented particularly low adherence rates among the poor and ethnic minorities within the U.S. (Cobaugh et al. 2008). Income inadequacy is a strong predictor for not filling prescription medications. One study that looked at racial disparities in the quality of medication use in older adults found that 28 percent of blacks could not purchase their medication due to cost compared to 12 percent of whites (Roth et al. 2009). Medication reconciliation post-discharge may be particularly important for patients with poor adherence to their medications, so the prescriber can evaluate what the patient is taking and reinforce which medications are most needed to improve their health.

Centers for Medicare and Medicaid Services. Racial, Ethnic, and Gender Disparities in Health Care in Medicare Advantage. 2019; 44, 1-137.

Cobaugh, D.J., E. Angner, C.I. Kiefe, M.N. Ray, C.L. Lacivita, N.W. Weissman, K.G. Saag, J.J. Allison. 2008. Effect of racial differences on ability to afford prescription medications. American Journal of Health-System Pharmacy: AJHP: Official Journal Of The American Society Of Health-System Pharmacists. 65, no. 22: 2137-43.

Roth, M. T., Esserman, D. A., Ivey, J. L., & Weinberger, M. Racial disparities in the quality of medication use in older adults: baseline findings from a longitudinal study. Journal of general internal medicine 2009; 25, 228-234.

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, Safety : Medication

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

Populations at Risk : Dual eligible beneficiaries, Populations at Risk : Individuals with multiple chronic conditions

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.)*

N/A

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

This is not an eMeasure **Attachment:**

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

Attachment**:** 0097_MRP_Fall_2020_Value_Sets.xlsx

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

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

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

Not an instrument-based measure

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

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.

This measure was updated to add Medicare as a product line and expand the age range to include Medicare patients 18 years and older. Prescription medication use is common among adults of all ages, particularly adults with chronic conditions, who make up the majority of Medicare Advantage (MA) patients. Expanding the measure to include all MA patients provides an opportunity to measure the quality of care coordination post-discharge, as well as patient safety.

The majority of patients served by SNPs and MA Plans are 65 and older; however, both can (and do) serve individuals who are under 65; for example, those with significant disability (e.g., dual-eligible). In 2013, 30 percent of hospital discharges in SNPs and 16 percent of hospital discharges in MA plans were for adults 18–64, who are as likely to benefit from medication reconciliation as older adults. Expanding the age range encourages medication reconciliation among all adults who are discharged from an inpatient facility.

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).

Medication reconciliation conducted by a prescribing practitioner, clinical pharmacist or registered nurse, as documented through either administrative data or medical record review on the date of discharge through 30 days after discharge (31 total days).

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).

Medication reconciliation conducted by a prescribing practitioner, clinical pharmacist or registered nurse, as documented through either administrative data or medical record review on the date of discharge through 30 days after discharge (31 total days). Medication reconciliation is defined as a type of review in which the discharge medications are reconciled with the most recent medication list in the outpatient medical record.

This measure is specified for medical record or administrative data collection.

Medical Record Reporting Details:

Documentation in the outpatient medical record must include evidence of medication reconciliation and the date when it was performed. Any of the following meets criteria:

- Documentation of the current medications with a notation that the provider reconciled the current and discharge medications.
- Documentation of the current medications with a notation that references the discharge medications (e.g., no changes in medications since discharge, same medications at discharge, discontinue all discharge medications).
- Documentation of the patient's current medications with a notation that the discharge medications were reviewed.
- Documentation of a current medication list, a discharge medication list and notation that both lists were reviewed on the same date of service.
- Documentation of the current medications with evidence that the patient was seen for post-discharge hospital follow-up with evidence of medication reconciliation or review. Evidence that the patient was seen for post-discharge hospital follow-up requires documentation that indicates the provider was aware of the patient's hospitalization or discharge.
- Documentation in the discharge summary that the discharge medications were reconciled with the most recent medication list in the outpatient medical record. There must be evidence that the discharge summary was filed in the outpatient chart on the date of discharge through 30 days after discharge (31 total days).
- Notation that no medications were prescribed or ordered upon discharge.

Only documentation in the outpatient medical record meets the intent of the measure, but an outpatient visit is not required.

Administrative Reporting Method Details:

See value sets provided for administrative codes meeting measure numerator intent.

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

All acute or nonacute inpatient discharges on or between January 1 and December 1 of the measurement year for patients who are 18 years and older.

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).

To identify an acute or nonacute inpatient discharge on or between January 1 and December 1 of the measurement year do the following:

1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set).
2. Identify the discharge date for the stay.

The denominator for this measure is based on discharges, not members. If members have more than one discharge, include all discharges on or between January 1 and December 1 of the measurement year.

If the discharge is followed by a readmission or direct transfer to an acute or nonacute inpatient care setting on the date of discharge through 30 days after discharge (31 total days), count only the last discharge. To identify readmissions and direct transfers during the 31-day period:

1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set).
2. Identify the admission date for the stay (the admission date must occur during the 31-day period).
3. Identify the discharge date for the stay (the discharge date is the event date).

Exclude both the initial and the readmission/direct transfer discharges if the last discharge occurs after December 1 of the measurement year.

If the admission date and the discharge date for an acute inpatient stay occur between the admission and discharge dates for a nonacute inpatient stay, include only the nonacute inpatient discharge. To identify acute inpatient discharges:

1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set).
2. Exclude nonacute inpatient stays (Nonacute Inpatient Stay Value Set).
3. Identify the admission date for the stay.
4. Identify the discharge date for the stay.

To identify nonacute inpatient discharges:

1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set).
2. Confirm the stay was for nonacute care based on the presence of a nonacute code (Nonacute Inpatient Stay Value Set).
3. Identify the admission date for the stay.
4. Identify the discharge date for the stay.

Additional guidance for identifying appropriate discharges for inclusion in the eligible population:

- If a patient remains in an acute or nonacute care setting through December 1 of the measurement year, a discharge is not included in the measure for this patient, but the organization must have a method for identifying the patient's status for the remainder of the measurement year, and may not assume the patient remained admitted based only on the absence of a discharge before December 1. If the organization is unable to confirm the patient remained in the acute or nonacute care setting through December 1, disregard the readmission or direct transfer and use the initial discharge date.

Additional guidance for identifying the eligible population:

Patients in hospice are removed from the eligible population.

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

No exclusions.

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.)*

N/A

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.)*

N/A

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

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

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

Better quality = Higher score

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

Step 1: Determine the eligible population. The eligible population is all the patients aged 18 years and older. Do not include patients who were discharged then subsequently readmitted to the hospital or directly transferred to another inpatient setting. Also do not include patients who received hospice services during the measurement year.

Step 2: Determine number of patients meeting the denominator criteria as specified in section S.9 above. The denominator includes all patients discharged from an inpatient facility. Patients may be counted more than once in the denominator if they had more than one discharge during the measurement year.

Step 3: Determine the number of patients who meet the numerator criteria as specified in section S.6 above. The numerator includes all patients who had a reconciliation of the discharge mediations with the current medication list in the outpatient medical record documented.

Step 4: Calculate the rate by dividing the total from Step 3 by the total from Step 2.

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.

This measure can be reported using administrative and/or medical record data. For organizations that choose to report the measure using medical record data, a systematic sample of 411 members is used. A sample size of 411 is used because it allows for the 95% confidence interval around the rate, meaning that a 5% difference in plan performance is statistically significant. Plans are instructed to list and sort all eligible members for a measure. NCQA then provides a Random Number table that organizations use to assist with sample selection. The Random Number table lists a value that is used to determine which members from the eligible population (i.e., every nth member) for whom numerator compliance will be determined.

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 Records, 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.

This measure is based on administrative claims and medical record documentation collected in the course of providing care to health plan patients. NCQA collects the Healthcare Effectiveness Data and Information Set (HEDIS) data for this measure directly from Health Maintenance Organizations and Preferred Provider Organizations via NCQA's online data submission system.

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

No data collection instrument provided

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

Health Plan

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

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.)*

N/A

2. Validity – See attached Measure Testing Submission Form

MRP_nqf_testing_attachment_7.1-637396684931453890.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)

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 0097
Measure Title: Medication Reconciliation Post-Discharge
Date of Submission: 11/2/2020

Type of Measure:

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

*cell intentionally left blank

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

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

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

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

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

2020 Submission

N/A

2015 Submission

N/A

1.3. What are the dates of the data used in testing?

2020 Submission

Testing of measure score reliability and validity was performed using HEDIS Health Plan performance data from January 1 to December 31, 2018.

2015 Submission

January 1 to December 31, 2013

1.4. What levels of analysis were tested? (*testing must be provided for* **all** *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

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	Measure Tested at Level of:
☒ health plan	☒ health plan
☐ other:	☐ other:

1.5. How many and which measured entities 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*)

2020 Submission

This measure assesses the number of discharges for Medicare members 18 years of age and older for whom medications were reconciled the date of discharge through 30 days after the discharge (31 total days). The intent of this measure is to prevent the inappropriate use of prescription medications. When patients are hospitalized, it puts them at higher risk for medication discontinuation, continuation or inadvertent starting of new medication which can impact existing conditions or other medications.

Sample for measure score reliability testing and construct validity testing: The measure score reliability was calculated from HEDIS data that included 472 Medicare health plans. The sample included all Medicare plans submitting data to NCQA for HEDIS. The plans were geographically diverse and varied in size.

2015 Submission

This measure was tested for reliability and meaningful difference in performance at the plan level using data from Medicare Special Needs Plans (SNP) submitting HEDIS data for the 2013 measurement year (HEDIS 2014). A total of 353 health plans that were nationally representative were included in this analysis. Special Needs Plans are Medicare Advantage health plans that provide care to dual eligible beneficiaries, individuals with chronic conditions and individuals residing in institutional settings.

1.6. How many and which patients 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*)

2020 Submission

HEDIS data are summarized at the health plan level for the Medicare product line. Below is a description of the sample. It includes number of health plans included in HEDIS data collection and the average and median eligible population for the measure across health plans.

Table 1. Mean and Median eligible population for *Medication Reconciliation Post-Discharge*, 2018

Product Type	Number of Plans	Mean number of eligible members per plan	Median number of eligible members per plan
Medicare	472	504	411

NOTE: Most health plans use a combination of data from administrative claims and a random sample of 411 medical records they review to report performance rates. However, there are some health plans that report

on the full population that qualifies for the denominator through administrative claims and there are some health plans that have fewer than 411 members who qualify for the denominator. This means that there is a range of denominator sizes reported to NCQA every year for this measure, but the median is generally 411.

Additionally, NCQA maintains detailed guidelines on the calculations and sampling that are used by health plans to report the measure, how to draw the sample of 411, guidance for oversampling when necessary, and how to handle denominators that are less than 411.

2015 Submission

In 2013, nearly 1.9 million Special Needs Plan beneficiaries were included in plans that reported HEDIS measures. Data is summarized at the health plan level for all Special Needs Plans submitting data for this measure for 2013. Patients included in the HEDIS data include a diverse representation of ages, race and diagnoses. The table below shows the average number of eligible patients per health plan and the standard deviation of that average across health plans.

Table 1: Sample 1 Average Eligible Population per Health Plan.

Product Type	Number of Plans	Average number of eligible patients per plan	Standard Deviation
Medicare SNP	353	656	1,247

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.

2020 Submission

No differences in the data used for reliability and construct validity testing.

2015 Submission

The data sample was used to demonstrate reliability (beta-binomial calculation), empirical validity (correlation analysis) and meaningful difference in performance.

Systematic evaluation of face validity: Validity was also demonstrated through a systematic assessment of face validity. This measure was systematically evaluated for face validity with three panels of experts: See Additional Information: Ad.1. Workgroup/Expert Panel Involved in Measure Development for names and affiliation of expert panels.

- The Geriatric MAP included 11 experts in geriatrics, including representation by consumers, health plans, health care providers and policy makers.
- The Technical Measurement Advisory Panel includes 12 members, including representation by health plans methodologists, clinicians and HEDIS auditors.
- NCQA's Committee on Performance Measurement (CPM) oversees the evolution of the measurement set and includes representation by purchasers, consumers, health plans, health care providers and policy makers. This panel is made up of 16 members. The CPM is organized and managed by NCQA and reports to the NCQA Board of Directors and is responsible for advising NCQA staff on the development and maintenance of performance measures. CPM members reflect the diversity of constituencies that

performance measurement serves; some bring other perspectives and additional expertise in quality management and the science of measurement.

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.

2020 Submission

We did not assess data by social risk factors. Social risk factor data were not available in reported results. This measure is specified for Medicare members age 18 and older. NCQA is actively engaged with partners including the CMS Office of Minority Health in identifying feasible methods to further integrate social risk factors into health plan quality measures, with a focus on stratification. This is aligned with recent recommendations from MedPAC and ASPE on optimal methods for addressing social risk in quality measurement and programs.[1,2] This is an NCQA wide initiative. Our intent is to implement methods to bridge data concerns in the future.

1. Medicare Payment Advisory Commission. (2020). The Medicare Advantage program: Status report. In Report to the Congress: Medicare Payment Policy (p. 397). http://medpac.gov/docs/default-source/reports/mar20_medpac_ch13_sec.pdf
2. Office of the Assistant Secretary for Planning and Evaluation, & U.S. Department of Health & Human Services. (2020). Second Report to Congress on Social Risk and Medicare's Value-Based Purchasing Programs. https://aspe.hhs.gov/social-risk-factors-and-medicares-value-basedpurchasing-programs

2015 Submission

N/A

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*)

2020 Submission

We utilized the methodology described by John Adams (Adams, J.L. The Reliability of Provider Profiling: A Tutorial. Santa Monica, California: RAND Corporation. TR-653-NCQA, 2009) to calculate signal-to-noise reliability. This methodology uses the Beta-binomial model to assess how well one can confidently distinguish the performance of one reporting entity from another. Conceptually, the Beta-binomial model is the ratio of signal to noise. The signal is the proportion of the variability in measured performance that can be explained by real differences across reporting entities (plans, physicians, etc.) in performance. The Beta-binomial model is an appropriate model when estimating the reliability of simple pass/fail rate measures, such as the Medication Reconciliation Post-Discharge measure. Reliability scores range from 0.0 to 1.0. A score of zero

implies that all variation is attributed to measurement error (i.e., noise), whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance across reporting entities.

For the *Medication Reconciliation Post-Discharge* measure, health plans are the reporting entity. For the formulas and explanations below, we use health plans as the reporting entity.

The formula for signal-to-noise reliability is:

Signal-to-noise reliability = $\sigma^2_{\text{plan-to-plan}} / (\sigma^2_{\text{plan-to-plan}} + \sigma^2_{\text{error}})$

Therefore, we need to estimate two variances: 1) variance between plans ($\sigma^2_{\text{plan-to-plan}}$); 2) variance within plans (σ^2_{error}).

1. Variance between plans = $\sigma^2_{\text{plan-to-plan}} = (\alpha\ \beta) / (\alpha + \beta + 1)(\alpha + \beta)^2$

   α and β are two shape parameters of the Beta-Binomial distribution, $\alpha > 0$, $\beta > 0$

2. Variance within plans: $\sigma^2_{\text{error}} = \hat{p}(1- \hat{p})/n$

   $\hat{p}$ = observed rate for the plan

   n = plan-specific denominator for the observed rate (most often the number of eligible plan members)

Using Adams' 2009 methodology, we estimated the reliability for each reporting entity, then averaged these reliability estimates across all reporting entities to produce a point estimate of signal-to-noise reliability. We label this point estimate "mean signal-to-noise reliability". The mean signal-to-noise reliability measures how well, on average, the measure can differentiate between reporting entity performance on the measure.

Along with the point estimate of mean signal-to-noise reliability, we are also providing:

1. The standard error (SE) and 95% confidence interval (95% CI) of the mean signal-to-noise reliability for all plans and stratified by the denominator size (number of eligible members per plan). The SE and 95% CI of the mean signal-to-noise reliability provides information about the stability of reliability. The 95% CI is the mean signal-to-noise reliability ± (1.96*SE). The narrower the confidence interval, the less the mean signal-to-noise reliability estimate will change due to idiosyncratic features of specific plans. We also stratified the results by the denominator size using terciles of the distribution to provide additional information about the stability of reliability.

2. The distribution (minimum, 10th, 25th, 50th, 75th, 90th, maximum) of the plan-level signal-to-noise reliability estimates. Each plan's reliability estimate is a ratio of signal to noise, as described above [$\sigma^2_{\text{plan-to-plan}} / (\sigma^2_{\text{plan-to-plan}} + \sigma^2_{\text{error}})$]. Variability between plans ($\sigma^2_{\text{plan-to-plan}}$) is the same for each plan, while the specific plan error (σ^2_{error}) varies. Reliability for each plan is an ordinal measure of how well one can determine where a plan lies in the distribution across plans, with higher estimates indicating better reliability. We also stratified the results by the denominator size, grouping plans with <411 in the denominator (Tercile 1), those with exactly 411 in the denominator (Tercile 2), and those with >411 in the denominator (Tercile 3), to provide additional information about the distribution of plan-level signal-to-noise reliability estimates. (For reasons unrelated to reliability, plans most often have denominators of exactly 411 for hybrid HEDIS measures.) The number of plans in each stratum and the per-plan denominators of the performance rates are displayed in the summary tables.

This methodology allows us to estimate the reliability for each plan and summarize the distribution of these estimates.

2015 Submission

Reliability Testing of Performance Measure Score: In order to assess measure precision in the context of the observed variability across accountable entities, we utilized the reliability estimate proposed by Adams (2009). The following is quoted from the tutorial which focused on provider-level assessment: "Reliability is a key metric of the suitability of a measure for [provider] profiling because it describes how well one can confidently distinguish the performance of one physician from another. Conceptually, it is the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in performance. There are three main drivers of reliability: sample size, differences between physicians, and measurement error. At the physician level, sample size can be increased by increasing the number of patients in the physician's data as well as increasing the number of measures per patient." This approach is also relevant to health plans and other accountable entities.

Adams' approach uses a Beta-binomial model to estimate reliability; this model provides a better fit when estimating the reliability of simple pass/fail rate measures as is the case with most HEDIS® measures. The beta-binomial approach accounts for the non-normal distribution of performance within and across accountable entities. Reliability scores vary from 0.0 to 1.0. A score of zero implies that all variation is attributed to measurement error (noise or the individual accountable entity variance) whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance (across accountable entities).

Adams, J. L. The Reliability of Provider Profiling: A Tutorial. Santa Monica, California: RAND Corporation. TR-653-NCQA, 2009

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*)

2020 Submission

Table 2a. Mean Signal-To-Noise Reliability, Standard Error (SE) and 95% Confidence Interval (95% CI) for the *Medication Reconciliation Post-Discharge* Measure by Terciles of the Denominator Size and for All Submissions, 2018

Stratification	Number of Plans	Number of Eligible Members per Plan (min max)	Mean Signal To Noise Reliability	SE	95% CI
All Medicare	**472**	**32 - 51839**	**0.977**	**0.001**	**(0.975, 0.979)**
Tercile 1	128	32 - 410	0.964	0.003	(0.959, 0.970)
Tercile 2	339	411 - 411	0.981	0.000	(0.981, 0.982)
Tercile 3	5	1442 - 51839	1.000	0.000	(0.999, 1.000)

SE: Standard Error of the mean.

95% CI: 95% confidence interval.

Table 2b. Distribution of Plan-Level Signal-To-Noise Reliability for the *Medication Reconciliation Post-Discharge* Measure by Terciles of the Denominator Size and for All Submissions, 2018

Stratification	Number of Plans	Distribution of Plan Estimates of Signal-to-Noise Reliability: Min	Distribution of Plan Estimates of Signal-to-Noise Reliability: P10	Distribution of Plan Estimates of Signal-to-Noise Reliability: P25	Distribution of Plan Estimates of Signal-to-Noise Reliability: P50	Distribution of Plan Estimates of Signal-to-Noise Reliability: P75	Distribution of Plan Estimates of Signal-to-Noise Reliability: P90	Distribution of Plan Estimates of Signal-to-Noise Reliability: Max
All Medicare	**472**	**0.804**	**0.961**	**0.981**	**0.982**	**0.986**	**0.990**	**1.000**
Tercile 1	128	0.832	0.922	0.955	0.976	0.984	0.988	1.000
Tercile 2	339	0.977	0.977	0.978	0.979	0.983	0.988	0.997
Tercile 3	5	0.999	0.999	1.000	1.000	1.000	1.000	1.000

2015 Submission

Results of reliability testing of performance measure score: The table below shows the results of the reliability testing of the performance measurement score in 2013.

# of plans	Overall Reliability Score	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
353	.98	.91	.95	.98	.99	.99

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?*)

2020 Submission

Table 2a provides the point estimate of mean signal-to-noise reliability, its standard error, and the 95% CI for the *Medication Reconciliation Post-Discharge* measure for Medicare plans overall and stratified by the denominator size (distribution of the number of eligible members per plan). Over all Medicare plans, the reliability estimate is 0.977, and the 95% CI is (0.975, 0.979), indicating very good reliability. Stratified analyses show that reliability increase as plan size gets larger.

Table 2b summarizes the distribution of plan-level signal-to-noise reliability estimates for the *Medication Reconciliation Post-Discharge* measure. Over all Medicare plans, the estimates range from 0.804 to 1.00. The 50th percentile is 0.982, which exceeds the 0.70 threshold for reliability. This table also includes the distribution of plan-level signal-to-noise reliability estimates stratified by denominator size. Reliability estimates are higher for plans with a larger denominator.

2015 Submission

Interpretation of measure score reliability testing:

Reliability scores can vary from 0.0 to 1.0. A score of zero implies that all variation is attributed to measurement error (noise) whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance (signal). Generally, a minimum reliability score of 0.7 is used to indicate sufficient signal strength to discriminate performance between accountable entities. The testing suggests that this measure has very good reliability. The 10-90th percentile distribution of health plan level-reliability on the rates in this measure

show all health plans met the minimally accepted threshold of 0.7, and the majority of plans exceeded 0.9. Strong reliability is demonstrated with the majority of variance attributed to signal and not to noise.

2b1. VALIDITY TESTING

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

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

☒ **Performance measure score**

☒ **Empirical validity testing**

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

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*
2020 Submission

Method of testing construct validity

We tested for construct validity by exploring the following:

- Is the *Medication Reconciliation Post-Discharge* measure correlated with the HEDIS *Transitions of Care measure indicators: Notification of Inpatient Admission, Receipt of Discharge Information* and *Patient Engagement After Inpatient Discharge?*

We hypothesized that three of the four rates within the *Transitions of Care* measure, specifically the *Notification of Inpatient Admission, Receipt of Discharge Information* and *Patient Engagement After Inpatient Discharge* indicators, would be highly positively correlated with the *Medication Reconciliation Post-Discharge* measure. Organizations that perform well on the *Medication Reconciliation Post-Discharge* measure should perform well on these indicators within the *Transitions of Care* measure, given that they address the same activities for patients transitioning from the inpatient to the outpatient setting.

NCQA performs Pearson correlation for construct validity using HEDIS health plan data. The test estimates the strength of linear association between two continuous variables; the magnitude of correlation ranges from -1 and +1. A value of 1 indicates a strong positive linear association: an increase in values of one variable is associated with increase in value of another variable. A value of 0 indicates no linear association. A value of -1 indicates a strong negative relationship in which an increase in values of the first variable is associated with a decrease in values of the second variable. The significance of a correlation coefficient is evaluated by testing the hypothesis that an observed coefficient calculated for the sample is different from zero. The resulting p-value indicates the probability of obtaining a difference at least as large as the one observed due to chance alone. We adjusted our p-values to account for testing multiple correlations and used a threshold of 0.05 to evaluate the test results. P-values less than this threshold imply that it is unlikely that a non-zero coefficient was observed due to chance alone.

Method of assessing face validity

NCQA develops measures using a standardized process. For new measures, face validity is assessed at various steps as described below.

STEP 1: NCQA staff identifies areas of interest or gaps in care. Clinical measurement advisory panels (MAPs), whose members are authorities on clinical priorities for measurement, participate in this process. Once topics are identified, a literature review is conducted to find supporting documentation on their importance, scientific soundness, and feasibility. This information is gathered into a work-up format, which is vetted by the MAPs, including the Geriatric Measurement Advisory Panel (GMAP), the Technical Measurement Advisory Panel (TMAP) and the Committee on Performance Measurement (CPM) as well as other panels as necessary.

STEP 2: Development ensures that measures are fully defined and tested before the organization collects them. MAPs participate in this process by helping identify the best measures for assessing health care performance in clinical areas identified in the topic selection phase. Development includes the following tasks: (1) Prepare a detailed conceptual and operational work-up that includes a testing proposal and (2) Collaborate with health plans to conduct field-tests that assess the feasibility and validity of potential measures. At this step, face validity is systematically determined by the CPM, which uses testing results and proposed final specifications to determine if the measure will move forward to Public Comment.

STEP 3: Public Comment is a 30-day period of review that allows interested parties to offer feedback to NCQA about proposed new measures. Public comment offers an opportunity to assess the validity, feasibility, importance and other attributes of a measure from a wider audience. For this measure, a majority of public comment respondents supported the measure. NCQA MAPs and the technical panels consider all comments and advise NCQA staff on appropriate recommendations brought to the CPM. Face validity is then again systematically assessed by the CPM. The CPM reviews all comments before making a final decision and votes to recommend approval of new measures for HEDIS. NCQA's Board of Directors then approves new measures.

This process was completed at the time of the measure's development and most recently during its reevaluation in 2015.

2015 Submission

Method of testing empirical validity: We tested for construct validity by exploring whether performance on Medication Reconciliation Post-Discharge was correlated with a similar measure of annual medication review by a prescribing provider (NQF #0553). We hypothesized that performance on these measures should be highly correlated because plans that excel at medication management should perform well on both measures. To test these correlations, we used a Person correlation test. This test estimates the strength of the linear association between two continuous variables: the magnitude of correlation ranges from -1 and +1. A value of 1 indicates a perfect linear dependence in which increasing values on one variable is associated with increasing values of the second variable. A value of 0 indicates no linear association. A value of -1 indicates a perfect linear relationship in which increasing values of the first variable is associated with decreasing values of the second variable.

Method of Assessing Face Validity: NCQA has identified and refined measure management into a standardized process called the HEDIS measure life cycle.

STEP 1: NCQA staff identifies areas of interest or gaps in care. Clinical expert panels (MAPs—whose members are authorities on clinical priorities for measurement) participate in this process. Once topics are identified, a literature review is conducted to find supporting documentation on their importance, scientific soundness and feasibility. This information is gathered into a work-up format. Refer to What Makes a Measure "Desirable"? The work-up is vetted by NCQA's Measurement Advisory Panels (MAPs), the Technical Measurement Advisory Panel (TMAP) and the Committee on Performance Measurement (CPM) as well as other panels as necessary.

STEP 2: Development ensures that measures are fully defined and tested before the organization collects them. MAPs participate in this process by helping identify the best measures for assessing health care performance in clinical areas identified in the topic selection phase. Development includes the following tasks: (1) Prepare a detailed conceptual and operational work-up that includes a testing proposal and (2) Collaborate with health plans to conduct field-tests that assess the feasibility and validity of potential measures. The CPM uses testing results and proposed final specifications to determine if the measure will move forward to Public Comment.

STEP 3: Public Comment is a 30-day period of review that allows interested parties to offer feedback to NCQA and the CPM about new measures or about changes to existing measures.
NCQA MAPs and technical panels consider all comments and advise NCQA staff on appropriate recommendations brought to the CPM. The CPM reviews all comments before making a final decision about Public Comment measures. New measures and changes to existing measures approved by the CPM will be included in the next HEDIS year and reported as first-year measures.

STEP 4: First-year data collection requires organizations to collect, be audited on and report these measures, but results are not publicly reported in the first year and are not included in NCQA's State of Health Care Quality, Quality Compass or in accreditation scoring. The first-year distinction guarantees that a measure can be effectively collected, reported and audited before it is used for public accountability or accreditation. This is not testing—the measure was already tested as part of its development—rather, it ensures that there are no unforeseen problems when the measure is implemented in the real world. NCQA's experience is that the first year of large-scale data collection often reveals unanticipated issues. After collection, reporting and auditing on a one-year introductory basis, NCQA conducts a detailed evaluation of first-year data. The CPM uses evaluation results to decide whether the measure should become publicly reportable or whether it needs further modifications.

STEP 5: Public reporting is based on the first-year measure evaluation results. If the measure is approved, it will be publicly reported and may be used for scoring in accreditation.

Step 6: Evaluation is the ongoing review of a measure's performance and recommendations for its modification or retirement. Every measure is reviewed for reevaluation at least every three years. NCQA staff continually monitors the performance of publicly reported measures. Statistical analysis, audit result review and user comments through NCQA's Policy Clarification Support portal contribute to measure refinement during re-evaluation. Information derived from analyzing the performance of existing measures is used to improve development of the next generation of measures.

Each year, NCQA prioritizes measures for re-evaluation and selected measures are researched for changes in clinical guidelines or in the health care delivery systems, and the results from previous years are analyzed. Measure workups are updated with new information gathered from the literature review, and the appropriate MAPs review the work-ups and the previous year's data. If necessary, the measure specification may be updated, or the measure may be recommended for retirement. The CPM reviews recommendations from the evaluation process and approves or rejects the recommendation. If approved, the change is included in the next year's HEDIS Volume 2.

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

2020 Submission

Table 3. Health-Plan Level Pearson Correlation Coefficients Among *Medication Reconciliation Post-Discharge* and the *Transitions of Care indicators, Notification of Inpatient Admission, Receipt of Discharge Information,* and *Patient Engagement After Inpatient Discharge* – **Medicare** Plans, calendar year 2018 data

*Significant at p<0.001

Measure	**Correlation Coefficient:** Notification of Inpatient Admission	**Correlation Coefficient:** Receipt of Discharge Information	**Correlation Coefficient:** Patient Engagement After Inpatient Discharge Rate
Medication Reconciliation Post-Discharge	0.45	0.43	0.60
(N=, p value =)	(449, p < 0.001)	(449, p < 0.001)	(472, p < 0.001)

Results of face validity assessment

Input from our multi-stakeholder measurement advisory panels (MAPs) and those submitting to public comment indicate the measure has good face validity—at both points—during the measure's development and reevaluation in 2015. Our MAPs agreed with the measure's intent and proposed specification, the majority of public comments received supported the measure, and our CPM, and subsequently our Board of Directors, approved the measure for HEDIS reporting.

2015 Submission

Results of empirical validity test:

The results in Table 1 indicated that for plan-level reporting this measure was significantly (p<.05) correlated with Medication Review (NQF #0553) in the direction that was hypothesized.

Table 1. Correlation between Medication Reconciliation Post-Discharge and Medication Review for Special Needs Plans – 2013

Pearson Correlation Coefficient Measure	Medication Reconciliation Post Discharge
Rate for the Care for Older Adults: Medication Review	R=0.4408 (R Statistic) p<.0001 (significance)

Note: All correlations are significant at p<.05

Results of face validity assessment:

Step 1: This measure was developed in 2007 to address the growing concern about medication errors that occur during transitions from the hospital to home. NCQA and the Geriatric MAP worked together to specify the measure.

Step 2: The measure was written and field-tested from 2007. After reviewing field test results, the CPM recommended to send the measure to public comment with a majority vote in 2008.

Step 3: The measure was released for Public Comment in 2008 prior to publication in HEDIS. We received and responded to 50 comments on this measure. The CPM recommended moving this measure to first year data collection by a majority vote.

Step 4: The measure was introduced in HEDIS 2009 and publicly reported for HEDIS 2011 after receiving a majority vote from the CPM.

Step 5: The measure is currently undergoing a re-evaluation which will go into HEDIS 2016.

Conclusion: The measure was deemed to have the desirable attributes of a HEDIS measure in 2009 (relevance, scientific soundness, and feasibility).

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?*)

2020 Submission

Interpretation of construct validity testing

Correlations between the *Medication Reconciliation Post-Discharge (MRP)* measure and the *Transitions of Care (TRC)* indicators, *Notification of Inpatient Admission, Receipt of Discharge Information* and *Patient Engagement After Inpatient Discharge,* were moderate to strong (Table 3). Plans with higher rates on *Medication Reconciliation Post-Discharge* tend to also have higher rates on the *Transitions of Care* indicators. The results indicate that the *Medication Reconciliation Post-Discharge (MRP)* measure has good validity.

Interpretation of systematic assessment of face validity

As stated above, the multi-stakeholder advisory panels, CPM and Board of Directors concluded, and public comment further indicated, the measure has good face validity.

2015 Submission

Interpretation of empirical validity testing: Coefficients with absolute value of less than 0.3 are generally considered indicative of weak associations whereas absolute values of 0.3 or higher denote moderate to strong associations. The significance of a correlation coefficient is evaluated by testing the hypothesis that an observed coefficient calculated for the sample is different from zero. The resulting p-value indicates the probability of obtaining a difference at least as large as the one observed due to chance alone. We used a threshold of 0.05 to evaluate the test results. P-values less than this threshold imply that it is unlikely that a non-zero coefficient was observed due to chance alone. The results confirmed the hypothesis that the Medication Reconciliation Post-Discharge measure and Medication Review measure are correlated with each other, suggesting that Medication Reconciliation Post-Discharge is a valid measure of a health plan's quality of medication management.

Interpretation of systematic assessment of face validity: These results indicate the technical expert panel showed good agreement that the measures as specified will accurately differentiate quality across providers. Our interpretation of these results is that this measure has sufficient face validity.

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions skip to section 2b3

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

N/A

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

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

_____

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO PM, or resource use measure, skip to section 2b4.
2020 Submission
N/A. Not an intermediate or health outcome, PRO-PM, or resource use measure.

2b3.1. What method of controlling for differences in case mix is used?
☐ **No risk adjustment or stratification**
☐ **Statistical risk model with risk factors**
☐ **Stratification by risk categories**
☐ **Other**

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

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)*

2020 Submission

To demonstrate meaningful differences in performance, NCQA calculates an inter-quartile range (IQR) for each indicator. The IQR provides a measure of the dispersion of performance. The IQR can be interpreted as the difference between the 25th and 75th percentile on a measure. To determine if this difference is statistically significant, NCQA calculates an independent sample t-test of the performance difference between two randomly selected plans at the 25th and 75th percentile. The t-test method calculates a testing statistic based on the sample size, performance rate, and standard error of each plan. The test statistic is then compared against a t-distribution, which is similar to a normal distribution. If the p value of the test statistic is less than 0.05, then the two plans' performance is significantly different from each other.

2015 Submission

To demonstrate meaningful differences in performance, NCQA calculates an inter-quartile range (IQR) for each indicator. The IQR provides a measure of the dispersion of performance. The IQR can be interpreted as the difference between the 25th and 75th percentile on a measure. To determine if this difference is statistically significant, NCQA calculates an independent sample t-test of the performance difference between two

randomly selected plans at the 25th and 75th percentile. The t-test method calculates a testing statistic based on the sample size, performance rate, and standardized error of each plan. The test statistic is then compared against a normal distribution. If the p value of the test statistic is less than .05, then the two plans' performance is significantly different from each other. Using this method, we compared the performance rates of two randomly selected plans, one plan in the 25th percentile and another plan in the 75th percentile of performance. We used these two plans as examples of measured entities. However, the method can be used for comparison of any two measured entities.

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*)

2020 Submission

Table 4a. Variation in Performance Across Health Plans, 2018

Measure	Avg. EP	Mean Rate	SD	10th	25th	50th	75th	90th	IQR
Medicare	504	0.61	0.18	0.36	0.50	0.63	0.75	0.84	0.25

EP: Eligible Population, the average denominator size across plans submitting to HEDIS

IQR: Interquartile range

Table 4b. T-Test

Measure	Plan Rate (25th Percentile)	Plan Rate (75th Percentile)	P Value
Medicare	0.50	0.75	$p < 0.001$

p-value: P-value of independent samples t-test comparing plans at the 25th percentile to plans at the 75th percentile.

2015 Submission

Variation in Performance across Health Plans in HEDIS (2013 data)

Measure	Avg. EP	Mean Rate	SD	10th	25th	50th	75th	90th	IQR
Medicare Special Needs Plans	319	36.6	21.1	9.4	19.2	34.7	52.8	62.1	33.6

EP: Eligible Population, the average denominator size across plans submitting to HEDIS

IQR: Interquartile range

T-test between two randomly selected health plans in HEDIS (2013 data)

Measure	Plan Rate (25th Percentile)	Plan Rate (75th Percentile)	P Value
Medicare Special Needs Plans	2.7	59.8	P<0.001

P-value: P-value of independent samples t-test comparing plans at the 25th percentile to plans at the 75th percentile

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?*)

2020 Submission

The results above indicate there is a 0.25 (25 percentage points) gap in performance between the 25th and 75th performing plans (see variation in performance across health plans table). The difference between the 25th and 75th percentile is statistically significant (see T-test table). This gap represents on average 126 more patients whose medications are reconciled in high performing Medicare plans compared to low performing plans (estimated from average health plan eligible population).

2015 Submission:

The results above indicate there is a 33.6% gap in performance between the 25th and 75th performing plans (see variation in performance across health plans table). The difference between the 25th and 75th percentile is statistically significant (see T-test table). This gap represents on average 107 more patients receiving medication reconciliation in high performing Medicare Special Needs Plans compared to low performing plans (estimated from average health plan eligible population).

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

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

2020 Submission

This measure has only one set of specifications.

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 non-responders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

2020 Submission

HEDIS measures apply to enrolled members in a health plan, and NCQA has a rigorous audit process to ensure the eligible population and numerator events for each measure are correctly identified and reported. The audit process is designed to verify primary data sources used to populate measures and ensure specifications are correctly implemented.

The HEDIS Compliance Audit addresses the following functions:

- Information practices and control procedures
- Sampling methods and procedures
- Data integrity
- Compliance with HEDIS specifications
- Analytic file production
- Reporting and documentation

2015 Submission

This measure is collected with a complete sample through medical record review, there is no missing data on this measure.

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*)

2020 Submission

HEDIS addresses missing data in a structured way through its audit process. HEDIS measures apply to enrolled members in a health plan, and NCQA-certified auditors use standard audit methodologies to assess whether data sources are missing data. If a data source is found to be missing data, and the issues cannot be rectified, the auditor will assign a "materially biased" designation to the measure for that reporting plan, and the rate will not be used. Once measures are added to HEDIS, NCQA conducts a first-year analysis to assess the measure's feasibility once widely implemented in the field. This analysis includes an assessment of how many plans report valid rates vs. rates that are materially biased. These considerations are weighed in the deliberation process before measures are approved for public reporting.

2015 Submission

This measure is collected with a complete sample through medical record review, there is no missing data on this measure.

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 non-responders) 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)

2020 Submission

This measure goes through the NCQA audit process each year to identify potential errors or bias in results. Only performances rates that have been reviewed and determined not to be "materially biased" are reported and used.

2015 Submission

This measure is collected with a complete sample through medical record review, there is no missing data on this measure.

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 by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition, Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

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

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

Some data elements are in defined fields in electronic sources

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

This measure is not currently specified as an eMeasure. However, health plans and providers that use an electronic health record to capture medication reconciliation use that data to report on this measure.

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.

NCQA conducts an independent audit of all HEDIS collection and reporting processes, as well as an audit of the data which are manipulated by those processes in order to verify that HEDIS specifications are met. NCQA has developed a precise, standardized methodology for verifying the integrity of HEDIS collection and calculation processes through a two-part program consisting of an overall information systems capabilities assessment followed by an evaluation of the managed care organization's ability to comply with HEDIS specifications. NCQA-certified auditors using standard audit methodologies will help enable purchasers to make more reliable comparisons between health plans.

The HEDIS Compliance Audit addresses the following functions:

1) Information practices and control procedures
2) Sampling methods and procedures
3) Data integrity
4) Compliance with HEDIS specifications
5) Analytic file production
6) Reporting and documentation

In addition to the HEDIS audit, NCQA provides a system to allow "real-time" feedback from measure users. Our Policy Clarification Support System receives thousands of inquiries each year on over 100 measures. Through this system, NCQA responds immediately to questions and identifies possible errors or inconsistencies in the implementation of the measure. This system informs both annual updates to the measures as well as routine re-evaluation of measures. These processes include updating value sets and clarifying the specifications. Measures are re-evaluated on a periodic basis and when there is a significant change in evidence.

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)*.

Broad public use and dissemination of these measures is encouraged and NCQA has agreed with NQF that noncommercial uses do not require the consent of the measure developer. Use by health care physicians in connection with their own practices is not commercial use. Commercial use of a measure requires the prior written consent of NCQA. As used herein, "commercial use" refers to any sale, license or distribution of a measure for commercial gain, or incorporation of a measure into any product or service that is sold, licensed or distributed for commercial gain, even if there is no actual charge for inclusion of the measure.

4. Usability and Use

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

4a. Accountability and Transparency

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

4.1. Current and Planned Use

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

Specific Plan for Use	Current Use (for current use provide URL)
*	**Public Reporting** **Annual State of Health Care Quality** **https://www.ncqa.org/report-cards/health-plans/state-of-health-care-quality-report/** **Health Plan Ratings** **http://www.ncqa.org/ReportCards/HealthPlans/HealthInsurancePlanRankings/HealthPlanRatingsPreview.aspx** **Annual State of Health Care Quality** **https://www.ncqa.org/report-cards/health-plans/state-of-health-care-quality-report/** **Health Plan Ratings** **http://www.ncqa.org/ReportCards/HealthPlans/HealthInsurancePlanRankings/HealthPlanRatingsPreview.aspx** **Payment Program** **Physician Quality Reporting Systems (PQRS)** **http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/** **Regulatory and Accreditation Programs** **HEDIS®-Health Plan** **http://www.ncqa.org/Programs/Accreditation/HealthPlanHP.aspx** **HEDIS®-Health Plan** **http://www.ncqa.org/Programs/Accreditation/HealthPlanHP.aspx** **Quality Improvement (external benchmarking to organizations)** **Quality Compass** **https://www.ncqa.org/programs/data-and-information-technology/data-purchase-and-licensing/quality-compass/** **CMS Medicare Advantage Plan Rating System (STARS)** **https://www.medicare.gov/find-a-plan/questions/home.aspx**

*cell intentionally left blank

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

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

HEALTH PLAN RATINGS/REPORT CARDS: This measure is used to calculate health plan rankings which are reported on the NCQA website. These rankings are based on performance on HEDIS measures among other factors. In 2019, a total of 255 Medicare health plans, 515 commercial health plans and 188 Medicaid health plans across 50 states were included in the rankings.

STATE OF HEALTH CARE ANNUAL REPORT: This measure is publicly reported nationally and by geographic regions in the NCQA State of Health Care annual report. This annual report published by NCQA summarizes findings on quality of care. In 2019, the report included results from calendar year 2018 for health plans covering a record 136 million people, or 43 percent of the U.S. population.
QUALITY COMPASS: This measure is used in Quality Compass which is an indispensable tool used for selecting a health plan, conducting competitor analysis, examining quality improvement and benchmarking plan performance. Provided in this tool is the ability to generate custom reports by selecting plans, measures, and benchmarks (averages and percentiles) for up to three trended years. Results in table and graph formats offer simple comparison of plans' performance against competitors or benchmarks.
HEALTH PLAN ACCREDITATION: This measure is used in scoring for accreditation of Medicare Advantage Health Plans. In 2019, a total of 247 Medicare Advantage health plans were accredited using this measure among others. Health plans are scored based on performance compared to benchmarks.
CMS MEDICARE ADVANTAGE PLAN RATING SYSTEM ("STARS"): This measure is included in the Medicare Advantage Star Rating System. CMS calculates a Star Rating (1-5) for all Medicare Advantage health plans based on 53 performance measures. Medicare beneficiaries can view the star rating and individual measure scores on the CMS Plan Compare website. The Star Rating is also used to calculate bonus payments to health plans with excellent performance. The Medicare Advantage Plan Rating program covers 11.5 million Medicare beneficiaries in 455 health plans across all 50 states.

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?*)
N/A

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.*)

N/A

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.

Health plans that report HEDIS calculate their rates and know their performance when submitting to NCQA. NCQA publicly reports rates across all plans and also creates benchmarks in order to help plans understand how they perform relative to other plans. Public reporting and benchmarking are effective quality improvement methods.

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.

NCQA publishes HEDIS results annually in our Quality Compass tool. NCQA also presents data at various conferences and webinars. For example, at the annual HEDIS Update and Best Practices Conference, NCQA presents results on measure performance. NCQA also regularly provides technical assistance on measures through its Policy Clarification Support System, as described in Section 3c1.

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.

NCQA measures are evaluated regularly. During this "reevaluation" process, we seek broad input on the measure, including input on performance and implementation experience. We use several methods to obtain

input, including vetting of the measure with several multi-stakeholder advisory panels, public comment posting, and review of questions submitted to the Policy Clarification Support System via MyNCQA.org. This information enables NCQA to comprehensively assess a measure's adherence to the HEDIS Desirable Attributes of Relevance, Scientific Soundness and Feasibility.

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

In general, health plans have not reported significant barriers to implementing this measure. Questions received through the Policy Clarification Support system have generally centered around minor clarification of the specifications, such as confirmation that information in claims or medical records meets the measure intent and satisfies the measure numerator and questions about the supporting guidelines for the measure. NCQA responded to all questions to ensure consistent implementation of the specifications. During the public comment period in 2015, a majority of comments from measured entities supported updates to the measure given its broad applicability and importance to all of Medicare patients discharged from the inpatient setting.

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

This measure has been deemed a priority measure by NCQA and other entities such as the Centers for Medicare and Medicaid Services as illustrated by its use in the programs listed above.

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.

During the measure's last major update, feedback obtained through the mechanisms described in 4a2.2.1 informed how we revised the measure specification including adding clarifying text and additional examples to further support determining numerator compliance.

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.

The 2016 to 2018 data shows that while performance rates for this measure are low, they have increased in the past year (see section 1b.2 for summary of data from health plans). In 2018, the average performance was 61.3%. There was a 47.7 percentage point difference between plans at the 10th and 90th percentiles. This large difference in performance represents a persistent gap in care and room for improvement in medication reconciliation for health plan members who use prescription medications. As noted in a prior section, 2016 average performance was 46.4 and 2018 average performance was 61.3, possibly reflecting revisions in the measure specification. The number of plans reporting has slightly increased over the years from 2016 (n=467) to 2018 (n=472). Overall performance has also increased over the years.

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.

No unintended consequences were identified during testing. No unintended consequences have been reported since this measure's implementation.

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

There were no identified unexpected findings during testing or since implementation of this measure.

5. Comparison to Related or Competing Measures

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

5. Relation to Other NQF-endorsed Measures

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

Yes

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

0419 : Documentation of Current Medications in the Medical Record

0553 : Care for Older Adults (COA) – Medication Review

2456 : Medication Reconciliation: Number of Unintentional Medication Discrepancies per Medication Per Patient

2988 : Medication Reconciliation for Patients Receiving Care at Dialysis Facilities

3317 : Medication Reconciliation on Admission

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?

Yes

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

See 5b.1 for more details.

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.)

This measure assesses medication reconciliation between a discharge medication list and a current medication list conducted post hospital discharge by a prescribing practitioner, clinical pharmacist or registered nurse and

documented in the outpatient record. The denominator for this measure is all discharges from an acute or nonacute facility for patients 18+.

Related Measures:

Measure 0553 is conducted at the Special Needs Plan (SNP) level. This measure assesses annual outpatient medication review (as distinct from reconciliation) by a prescribing practitioner or clinical pharmacist among patients aged 66+. A hospital discharge is not required to meet denominator criteria therefore the measure has a different target population than measure 0097 and is not a competing measure.

Measure 2456 is conducted at the hospital/acute facility level. This measure assesses the quality of the medication reconciliation process in the hospital by identifying errors in admission and discharge medication orders due to problems with the medication reconciliation process. This process is completed by a trained pharmacist who at the time of admission, compares the admission orders to the preadmission medication list to look for discrepancies and identify which discrepancies were unintentional using brief medical record review. This measure does not address whether a reconciled medication list is documented in the outpatient medical record after discharge. Therefore the measure focus is different from measure 0097.

Measure 0419e is conducted at the provider level. This measure looks at the percentage of visits for all patients 18+ for which the eligible professional attests to documenting a list of current medications using all immediate resources available on the date of the encounter. The list must include all known prescriptions, over-the-counters, herbals, and vitamin/mineral/dietary supplements AND must contain the medications' name, dosage, frequency and route of administration. This measure only looks for documentation of current medications and is not focused on reconciling medications after a discharge. The measure has a different target population and measure focus and is therefore not competing.

Measure 3317 is conducted at the facility level. This measure assesses the percentage of patients for whom a designated prior to admission (PTA) medication list was generated by referencing one or more external sources of PTA medications and for which all PTA medications have a documented reconciliation action by the end of Day 2 of the hospitalization. The list may include prescriptions, over-the-counter medications, herbals, vitamin/mineral/dietary (nutritional) supplements, and/or medical marijuana. This measure only looks at whether the medication should be continued, discontinued or modified. Given this measure targets medications prior to an admission and assesses adult and pediatric patients it is not competing.

Measure 2988 is conducted at the facility level. This measure assesses the percentage of patient-months for which medication reconciliation was performed and documented by an eligible professional. All known home medications (prescriptions, over-the-counters, herbals, vitamin/mineral/dietary (nutritional) supplements, and medical marijuana) need to be reconciled. The target population is members receiving dialysis and the measure aims to assess the use of at-home medications and compare them with medications in the dialysis medical record. This measure is different because of the target population and focus and therefore is not competing.

Appendix

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

No appendix **Attachment:**

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): National Committee for Quality Assurance

Co.2 Point of Contact: Bob, Rehm, nqf@ncqa.org, 202-955-1728-

Co.3 Measure Developer if different from Measure Steward: National Committee for Quality Assurance

Co.4 Point of Contact: Brittany, Wade, wade@ncqa.org, 202-530-0463-

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.

Geriatric Measurement Advisory Panel (GMAP):

Wade Aubry, MD, University of California, San Francisco

Arlene Bierman, MD,MS, Agency for Healthcare Research and Quality (AHRQ)

Patricia Bomba, MD, MACP, FRCP, Excellus BlueCross BlueShield

Nicole Brandt, PharmD, MBA, BCPP, BCGP, FASCP, University of Maryland, School of Pharmacy

Jennie Chin Hansen, RN, MS, FAAN, Geriatric Expert

Joyce Dubow, MUP, Consumer Advocate

Peter Hollmann, MD, Brown University

Jeffrey Kelman, MD, CMS

Karen Nichols, MD, Trinity-Health PACE

Steven Phillips, MD, CMD, Geriatric Specialty Care

Erwin Tan, MD, American Association of Retired Persons (AARP)

Eric G Tangalos, MD, Mayo Clinic

Dirk Wales, MD, PsyD, Axial Healthcare

Joan Weiss, PhD, RN, CRNP, FAAN, Health Resources and Services Administration

Neil Wenger, MD, MPH, UCLA Division of General Internal Medicine and RAND

Committee on Performance Measurement (CPM):

Andrew Baskin, MD, CVS Health/Aetna

Elizabeth Drye, MD, SM Yale School of Medicine

Mark Friedberg, MD, MPP, Blue Cross Blue Shield Massachusetts

Andrea Gelzer, MD, MS, FACP, AmeriHealth Caritas

David Grossman, MD, MPH, Washington Permanente Medical Group

Christine Hunter, (Co-Chair), MD, Independent Board Director

David Kelley, MD, MPA, Pennsylvania Department of Human Services

Jeffrey Kelman, MMSc, MD, CMS

Nancy Lane, PhD, Independent Consultant

Bernadette Loftus, MD, Perspicax Healthcare

Amanda Parsons, MD, MetroPlus

Wayne Rawlins, MD, MBA, ConnectiCare

Misty Roberts, MSN, RN, CPHQ, Humana

Rudy Saenz, MD, MMM, FACOG, Riverside Medical Clinic

Marcus Thygeson, (Co-Chair) MD, MPH, Blind Benefits

JoAnn Volk, MA, Georgetown University

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2009

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

Ad.4 What is your frequency for review/update of this measure? Measures are re-evaluated on a periodic basis and when there is a significant change in evidence.

Ad.5 When is the next scheduled review/update for this measure? 12, 2021

Ad.7 Disclaimers: These performance measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Ad.8 Additional Information/Comments: NCQA Notice of Use. Broad public use and dissemination of these measures is encouraged and NCQA has agreed with NQF that noncommercial uses do not require the consent of the measure developer. Use by health care physicians in connection with their own practices is not commercial use. Commercial use of a measure requires the prior written consent of NCQA. As used herein, "commercial use" refers to any sale, license or distribution of a measure for commercial gain, or incorporation of a measure into any product or service that is sold, licensed or distributed for commercial gain, even if there is no actual charge for inclusion of the measure.