All Cause Admissions and Readmissions, Spring 2019 Review Cycle: CDP Report

TECHNICAL REPORT

February 21, 2020

NATIONAL QUALITY FORUM

This report is funded by the Department of Health and Human Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001.

Contents

All Cause Admissions and Readmissions, Spring 2019 Review Cycle

TECHNICAL REPORT

Executive Summary

Avoidable hospital admissions and readmissions are an important focus for healthcare quality improvement. These avoidable admissions and readmissions often represent an opportunity to improve patient care transitions and prevent the unnecessary exposure of patients to adverse events in an acute care setting. NQF currently has 51 endorsed All Cause and condition-specific admissions and readmissions measures for various settings. Several federal quality improvement programs have adopted these measures to reduce unnecessary admissions and readmissions to improve communication and care transitions.

For this project, the Standing Committee evaluated two newly submitted measures and one measure undergoing maintenance review against NQF's standard evaluation criteria. Newly submitted measure 3495 *Hospital-Wide 30-Day, All Cause, Unplanned Readmission Rate (HWR) for the Merit-Based Incentive Payment System (MIPS) Eligible Clinician Groups* was originally submitted as one measure but was later split into two levels of analysis (Clinician Group/Practice and Individual Clinician) and evaluated separately, due to concerns related to attribution. The Committee also considered NQF 2539 *Facility 7-Day Risk Standardized Hospital Visit Rate after Outpatient Colonoscopy*.

The following measures are not endorsed:

- 3495 Hospital-Wide 30-Day, All Cause, Unplanned Readmission Rate (HWR) for the Merit-Based Incentive Payment System (MIPS) - *Individual clinician level of analysis*
- 3443 All Cause Emergency Department Utilization Rate for Medicaid Beneficiaries with Complex Care Needs and High Costs (BCNs)
- 3445 All Cause Inpatient Admission Rate for Medicaid Beneficiaries with Complex Care Needs and High Costs (BCNs)

The following measure was deferred to the Fall 2019 cycle:

- 3495 Hospital-Wide 30-Day, All Cause, Unplanned Readmission Rate (HWR) for the Merit-Based Incentive Payment System (MIPS): Clinician: Groups level of analysis

The following measure was withdrawn during the Committee' review:

- 2539 Facility 7-Day Risk Standardized Hospital Visit Rate after Outpatient Colonoscopy.

Measure 3495 *Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate for the Merit-Based Incentive Payment System (MIPS)* was sent back to the Standing Committee for reconsideration at the Clinician Group/Practice level of analysis only.

Throughout the Spring 2019 cycle, the Committee experienced challenges with achieving quorum during the measure evaluation webinars, leading to voting via survey after the call as per NQF's standard process. Additionally, during the Committee's post-comment call on October 2, 2019, in response to comments received during the public comment period and questions raised by the Committee concerning reliability testing, the developer inadvertently reported incorrect measure score reliability results during the live call, potentially influencing the Committee's deliberation and eventual decision to vote to reconsider their recommendation for endorsement of 3495 *Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate for the Merit-Based Incentive Payment System (MIPS)* at the Clinician Group/Practice level of analysis. After the post-comment call, the developer provided the correct reliability results for the minimum case number of 25 patients in writing, which had also been provided in their original submission. In consultation with the developers and the Committee co-chairs, it was determined to return measure 3495 to the Standing Committee for re-evaluation at the Clinician Group/Practice level of analysis, during the Fall 2019 cycle, due to concerns that the incorrect information provided may have influenced the Committee's deliberations and vote.

Measure 2539 *Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy* was withdrawn from consideration pending alignment of measure testing and specifications.

The Consensus Standards Approval Committee upheld the All Cause Admissions and Readmissions Standing Committee's recommendation not to endorse 3495 *Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate* at the Individual Clinician level of analysis and upheld the recommendation to send 3495 *Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate* at the Group Clinician/Practice level of analysis back to the Standing Committee.

Additionally, two measures from the Fall 2018 review cycle are included in this report. NQF 3443 *All Cause Emergency Department Utilization Rate for Medicaid beneficiaries with Complex Care Needs and High Costs (BCNs)* and NQF 3445 *All Cause Inpatient Admission rate for Medicaid Beneficiaries with Complex Care Needs and High Costs (BCNs)*. During the Fall 2018 cycle, the Readmissions Committee did not recommend these measures for endorsement.

The measure developer for measures 3443 and 3445 submitted requests for reconsideration to the Committee; the Committee reconsidered the developer's requests during the Spring 2019 review cycle and ultimately decided to uphold its initial decision not to recommend the measures for endorsement.

Brief summaries of the measures currently under review are included in the body of the report; detailed summaries of the Committee's discussion and ratings of the criteria for each measure are in Appendix A.

Introduction

Avoidable admissions and readmissions to acute care facilities are an important area for healthcare quality improvement. These avoidable admissions and readmissions often represent an opportunity to improve care transitions and prevent the unnecessary exposure of patients to adverse events in an acute care setting. To drive improvement in admissions and readmissions, performance measures have continued to be a key element of value-based purchasing programs to incentivize collaboration in the healthcare delivery system. Shared accountability is required to improve this health outcome, as many healthcare providers have a role in ensuring a safe patient transition between care settings. While a wide variety of healthcare stakeholders support the goal of reducing unnecessary hospitalizations, debates remain on the target rate of readmissions, appropriate methods for attribution, and if these performance measures should be linked to provider payment.

While admissions and readmissions are important patient outcomes, systematic reviews have found that less than a third of readmissions are preventable.[1] Many factors influence the rate of admissions and readmissions, including the resources available in the community to support a safe transition between care settings and the social support available to patients. While these factors have a role, poor care coordination and low-quality care also led to higher rates of readmission. Evidence demonstrates that provider interventions can improve these important patient outcomes, such as improved communication of patient discharge instructions, coordination with post-acute care providers and primary care physicians, and the reduction of complications such as hospital-acquired conditions.[2–4]

To incentivize reductions in inappropriate hospitalizations, CMS expanded accountability for avoidable readmissions throughout its quality reporting and payment programs. The Hospital Readmissions Reduction (HRRP) program reduces payment rates to hospitals with higher-than-expected readmission rates. The Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act) required CMS to implement quality measures for potentially preventable readmissions to long-term care hospitals, inpatient rehabilitation facilities, skilled nursing facilities, and home health agencies. Finally, CMS' Merit-Based Incentive Payment System (MIPS), which adjusts Medicare payments at the physician level, includes an option of an All Cause hospital readmission measure for groups with at least 16 clinicians and a sufficient number of cases.[5] Groups that report on the readmission measure are eligible for higher payment rates than clinician groups that do not. Given the increased use of readmission measures across settings of care, ensuring their scientific merit is more important than ever.

In this project, the All Cause Admissions and Readmissions Standing Committee considered NQF 3495 *Hospital Wide Unplanned Readmission* for endorsement; this measure is split into two levels of analysis – clinician group/practice and individual clinician level of analysis. Each level of analysis was assessed separately. The Committee also considered NQF 2539 *Facility 7-Day Risk Standardized Hospital Visit Rate after Outpatient Colonoscopy*.

Additionally, the Committee reviewed the reconsideration requests from the measure developer for NQF 3443 and NQF *3445 All Cause Inpatient Admission Rate for Medicaid Beneficiaries with complex care needs and high costs (BCNS).* Measures 3443 and 3445 were not recommended for endorsement in the Fall 2018 review cycle and the developer requested reconsideration during this cycle. This request was discussed by the Committee during the spring 2019 review cycle and ultimately the Committee decided not to reconsider either measure and upheld its initial decision not to recommend the measures.

NQF Portfolio of Measures for All Cause and Condition-Specific Admissions and Readmissions

The All Cause Admissions and Readmissions Standing Committee (Appendix C) oversees NQF's portfolio of admissions and readmissions measures (Appendix B) that includes measures for a number of different sites of care. This portfolio contains 51 measures:

Table 1. NQF Admissions and Readmissions Portfolio of Measures

	All Cause	Condition-Specific
Hospital	5	14
Home health	4	0
Skilled nursing facility	4	0
Long-term care facility	1	0
Inpatient rehab facility	1	0
Inpatient psychiatric facility	1	0
Dialysis facility	2	0
Health plan	1	0
Population-based	4	11
Hospital outpatient/ambulatory surgery center	0	1
Integrated delivery system	1	0
Accountable care organizations (ACO)	1	0
Total	25	26

Additional measures are assigned to other portfolios. These include patient-reported outcome and transition-of-care measures (Patient Experience and Function), and a variety of condition-specific readmission measures (Surgery and Perinatal).

Readmissions Measure Evaluation

On June 20 and 21, 2019, the Readmissions Standing Committee evaluated one new measure and one measure undergoing maintenance review against NQF's standard measure evaluation criteria. During this time, the Committee also reviewed the request for reconsideration for the two measures from the Fall 2018 cycle and voted not to take up the reconsideration.

Table 2. All Cause Admissions and Readmissions Measure Evaluation Summary

	Maintenance	New	Total
Measures under consideration	1	4*	5
Measures deferred	0	1	1
Measures not recommended for endorsement	0	3**	3
Measures withdrawn from consideration	1	0	1

*includes one measure reviewed as two separate measures based on different a level of analysis
**includes two measures with a developer request for reconsideration from the Fall 2018 review cycle

Comments Received Prior to Committee Evaluation

NQF solicits comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). In addition, NQF solicits comments for a continuous 16-week period during each evaluation cycle via an online tool located on the project webpage. For this evaluation cycle, the commenting period opened on May 1, 2019 and closed on August 30, 2019. As of June 12, 2019, two comments were submitted and shared with the Committee prior to the measure evaluation meetings (Appendix D).

All submitted comments were provided to the Committee prior to its initial deliberations during the Committee webinars.

Comments Received After Committee Evaluation

The continuous 16-week public commenting period with NQF member support closed on August 30, 2019. Following the Committee's evaluation of the measures under review, NQF received one comment from one-member organization pertaining to the draft report and to the measures under consideration. All comments for each measure under consideration have been summarized in Appendix A.

Throughout the 16-week continuous public commenting period, NQF members had the opportunity to express their support ('support' or 'do not support') for each measure submitted for endorsement consideration to inform the Committee's recommendations. One member was not supportive of measures 3495 (at both levels of analysis) and 2539.

Overarching Issues

During the Standing Committee's discussion of the measures, several overarching issues emerged that were factored into the Committee's ratings and recommendations for multiple measures.

Shared Accountability

Improvement of avoidable admissions and readmissions requires shared accountability among multiple healthcare providers. In this project, the Committee reviewed two readmissions measures that attribute the outcome to individual physicians, physician groups, along with hospital outpatient departments (HOPD) or ambulatory surgical centers (ASC) following a colonoscopy procedure. While the Committee agreed that these providers and settings have a role, the readmission of one patient can be counted in several measures, assessing quality for multiple providers. For example, a readmission could count in the numerator of a measure assessing a hospital's readmission rate as well as a physician group's readmissions performance rate.

While shared accountability is required to improve this outcome, the Committee struggled with attribution to an individual clinician. Attribution of readmission rates should consider the locus of control of the accountable entity. While multiple factors influence this outcome, individual physicians often rely on peers, office and hospital-based support staff for important care transition roles and may have limited direct influence on the social risk and the community social supports available to their patients. The Committee noted that shared accountability should be balanced with the locus of control of the accountable unit.

Social Risk

The use of readmission measures for payment has raised questions about how much control a healthcare provider can have over a patient's outcomes, as healthcare outcomes are influenced by both the care received and patient factors. In particular, stakeholders have raised concerns about the potential impact of social risk factors, as there is growing evidence demonstrating how these factors can influence health outcomes. The Committee recognized this evidence and reiterated the need to consider the potential influence of social risk factors on the results of admission and readmission measures. The Committee noted the need to ensure that healthcare providers disproportionately serving communities with increased social risk factors are not penalized unfairly, especially when readmission measures are publicly reported or used to determine payment. The Committee emphasized the need to maximize the predictive value of a risk-adjustment model and noted its expectation that developers will continue testing the risk-adjustment model with additional social risk factors to better understand unmeasured patient risk. The Committee noted that the conceptual rationale for adjusting for social risk should be considered on a case-by-case basis.

Unintended Consequences

The Committee reiterated the concern that readmission measures should account for a potential increase in observation stays and emergency department holding as an unintended negative consequence to patients. Some argue that patients may prefer treatment in these settings if possible,[6] while others note that patients may experience negative consequences from observation stays such as less timely and less coordinated care.[7] Observation stays can occur in the emergency department, in a dedicated unit, or in a setting similar to being admitted as an inpatient, leading to varying patient experience and time in the hospital.[8] Finally, patients may incur financial hardship if they require post-acute care after an observation stay, as Medicare will not cover a skilled nursing facility stay after an observation stay.[9] Because of the potential consequences to patients, the Committee recognized the

need to continue to monitor for increased use of emergency department (ED) visits and observation stays as potential consequences of the use of readmission measures.

Summary of Measure Evaluation

The following brief summaries of the measure evaluation highlight the major issues that the Committee considered. Details of the Committee's discussion and ratings of the criteria for each measure are included in Appendix A.

3495 Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate for the Merit-Based Incentive Payment System (Centers for Medicare & Medicaid Services/Yale New Haven Health Services Corporation Center for Outcomes Research and Evaluation): Not Recommended at the Individual Clinician Level of Analysis; Deferred to Fall 2019 at the Group Clinician Level of Analysis

Description: This measure is a re-specified version of the hospital-level measure, "Hospital-Wide All Cause, Unplanned Readmission Measure" (NQF 1789), which was developed for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare and are hospitalized in non-federal hospitals. This re-specified measure attributes hospital-wide index admissions to up to three participating MIPS Eligible Clinicians or Eligible Clinician Groups ("providers"), rather than to hospitals. It assesses each provider's rate of 30-day readmission, which is defined as unplanned, All Cause readmission within 30 days of hospital discharge for any eligible condition. The measure reports a single summary risk adjusted readmission rate (RARR), derived from the volume-weighted results of five different models, one for each of the following specialty cohorts based on groups of discharge condition categories or procedure categories: surgery/gynecology; general medicine; cardiorespiratory; cardiovascular; and neurology.
Measure Type: Outcome; **Level of Analysis**: Clinician: Group/Practice, Clinician: Individual; **Setting of Care**: Inpatient/Hospital; **Data Source**: Claims, Enrollment Data

INDIVIDUAL CLINICIAN LEVEL OF ANALYSIS

The Standing Committee reviewed the distribution of risk-adjusted readmission ratios (RARRs) for eligible clinicians' ranges from a mean of 11.4 in the first decile to 19.5 in the tenth decile. Similar to the review of the physician-group specification of this measure, the Standing Committee reviewed the input provided by the NQF Scientific Methods Panel. The Standing Committee generally agreed with the SMP subgroup members that the approach to reliability testing was appropriate. The Standing Committee reviewed the validity testing provided by the developer, the risk-adjustment methodology, and the calibration statistics.

The Standing Committee raised several concerns about the individual clinician specifications and testing. Specifically, the Committee was concerned about the attribution approach. The face validity of holding an individual clinician accountable for All Cause hospital readmission raised concern. The Committee was not clear how an individual physician could directly influence these outcomes without collaboration of other physicians and hospital partnerships. There was general agreement that the measure uses claims data that can be operationalized; however, the measure is not yet in use. There are no fees, licensing, or requirements to use the measure. The Standing Committee acknowledged that this measure is planned for use in the CMS MIPS program. The Standing Committee had concerns about the usability of a

clinician-level readmission measure and the extent to which it would provide the information necessary to implement targeted quality improvement efforts. The Standing Committee noted that the benefits of the performance measure may not outweigh evidence of unintended negative consequences to individuals or populations given the limited ability of an individual clinician to influence the outcome. The Standing Committee generally agreed that the clinician level of analysis of this measure does not meet the NQF criteria for endorsement.

CLINICIAN: GROUP/PRACTICE LEVEL OF ANALYSIS

The Standing Committee reviewed the logic model presented by the developer demonstrating physician group interventions that can reduce the risk of unplanned hospital visits. Several Standing Committee members agreed that physician groups and ACOs should have the infrastructure to improve admissions and readmissions.

The Standing Committee reviewed the input provided by the NQF Scientific Methods Panel (SMP). The Standing Committee noted the SMP's concerns that social risk factors are excluded from the risk model and the potential for negative consequences on access to care if this measure is not adequately risk adjusted.

The Standing Committee agreed that the developer should examine other clinical variables such as frailty or functional status. The Standing Committee reviewed the measure score empirical validity testing and face validity testing submitted at the individual physician and physician group levels. The Standing Committee noted that the evidence and validity testing must be evaluated separately for the two levels of analysis. The Committee agreed to vote individually on the two levels of analysis. The Standing Committee noted that eligible clinician groups' risk-adjusted readmission rates go down with increasing Overall Hospital Quality Star Rating and with increasing quintile of the Star Rating readmission quality score. The Committee reviewed the face validity testing and results, the approach to risk adjustment, and the conceptual model for sociodemographic risk adjustment. Committee members had differing views of the face validity of the measure with respect to the role that physician groups have in improving this outcome.

The Standing Committee agreed that the measure uses claims data that can be operationalized; however, the measure is not yet in use. There are no fees, licensing, or requirements to use the measure. The Standing Committee acknowledged that this measure is planned for use in the CMS MIPS program. The Standing Committee noted that this is a new measure and there is no information available on performance improvement. At the time of their original discussion, the Standing Committee agreed that the clinician group level of analysis of this measure generally met the NQF criteria of endorsement.

One comment was received on this measure at both levels of analysis, raising both process concerns and measure-specific concerns. The process concerns noted the lack of quorum during the Committee measure evaluation webinar on June 21 and the posting of a draft report that omitted vote counts. The commenter raised several concerns with the measure's evidence; the assignment of responsibility to

multiple physicians and practices; the measure's reliability at the minimum case number of 25 patients; and the conceptual basis used to explain which social risk factors were tested. The Committee discussed the comment extensively on the post-comment call. Additionally, during the post-comment call, the developer inadvertently reported incorrect measure score reliability results verbally, influencing the Committee's deliberation. The Committee ultimately voted not to reconsider the measure and to continue to recommend the measure at the clinician: group level.

Following the post-comment call, the developer submitted additional reliability testing information and analyses at various case volumes to respond to the Committee questions. Further, the developer noted that incorrect reliability testing information for the minimum case number of 25 patients had been given verbally during the post-comment call, and the correct information was in the submission form provided. Given that this incorrect information informed the Committee's decision, NQF, in consultation with the Committee co-chairs and the measure developers and stewards, decided to defer the final endorsement decision on this measure. This measure will go back to the Committee in the Fall 2019 cycle for re-review at the Clinician Group level of analysis, and a final endorsement recommendation will be made during that cycle.

Measure Withdrawn from Consideration

2539 Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy (Centers for Medicare & Medicaid Services/Yale New Haven Health Services Corporation Center for Outcomes Research and Evaluation): Withdrawn following Committee Discussion

Description: Facility-level risk-standardized rate of acute, unplanned hospital visits within 7 days of a colonoscopy procedure performed at a hospital outpatient department (HOPD) or ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older. An unplanned hospital visit is defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission. The measure is calculated separately for ASCs, and HOPDs. **Measure Type**: Outcome; **Level of Analysis**: Facility; **Setting of Care**: Outpatient Services; **Data Source**: Claims, Other

The Standing Committee reviewed this measure of hospital visits within 7 days of a colonoscopy procedure performed at a hospital outpatient department (HOPD) or ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older. The measure captures adverse patient outcomes associated with HOPD and ASC care, an important area for quality improvement. The Standing Committee agreed that the measure provides a logic model demonstrating provider-level and facility-level interventions that can to reduce the risk of unplanned hospital visits. These provider-level factors include a protocol for a patient's colonoscopy prep and technical quality of the procedure. The facility-level factors include anesthesia, discharge, and follow-up protocols.

The Standing Committee reviewed the overall 25th to 75th percentile performance range of risk-standardized hospital visit rates per 1000 colonoscopies from 11.8 percent to 12.8 percent, with mean performance of 12.3 percent and generally agreed that there is a narrow performance gap. The Standing Committee reviewed the reliability testing and noted differences in the measure specifications (one year

of data) and the three years of simulated signal-to-noise testing data. The Standing Committee requested that the developer consider aligning the testing with the measure specifications and resubmitting the measure using a three-year time frame for consideration in a future review. The Standing Committee noted that this alignment would help facilitate transparency and understanding among stakeholders and those being measured. The developer agreed to withdraw the measure from consideration and resubmit in a future cycle with testing data and measure specifications using a three-year time frame. Given the withdrawal from the process, voting on the measure was suspended.

The developer withdrew the measure from consideration and will update the measure and submit it to NQF in a future measure review cycle.

Requests for Reconsideration

3443 All Cause Emergency Department Utilization Rate for Medicaid Beneficiaries with Complex Care Needs and High Costs (BCNs) (Mathematica Policy Research): Not Recommended

Description: All Cause emergency department (ED) utilization rate for adult Medicaid beneficiaries who meet BCN population eligibility criteria. The measure is calculated as the number of ED visits per 1,000 beneficiary months and is intended to be reported at the state level. For the purpose of this measure, the BCN population is defined as Medicaid beneficiaries who are age 18 to 64 during the lookback year (the 12 months prior to the measurement year) and the measurement year and have at least one inpatient admission and at least two chronic conditions, as defined by the Chronic Conditions Data Warehouse (CCW), during the lookback year. Beneficiaries dually enrolled in Medicaid and Medicare and beneficiaries who had fewer than 10 months of Medicaid eligibility in the lookback year are not included in the analytic sample because we did not have enough utilization data to include them in testing. We further limited the analytic file to beneficiaries that met the BCN definition criteria described above. **Measure Type**: Outcome; **Level of Analysis**: Population: Regional and State; **Setting of Care**: Emergency Department and Services; **Data Source**: Claims

NQF 3443 assesses All Cause emergency department utilization for Medicaid beneficiaries between 18 and 64 years old who meet the criteria of presenting complex care needs and high costs. This measure was developed with the intention of pairing it with NQF 3445. The Committee agreed there was sufficient evidence that the measured entity could influence the outcome. Specifically, the Committee noted that the developer cited several studies demonstrating that emergency department visits in complex patients could be reduced through improved care management and agreed that performance varied. The developer conducted signal-to-noise (SNR) reliability testing for this measure using MAX data from 10 states. Committee members did not have any concerns about the reliability of the measure. However, the Committee raised a number of points under the validity subcriterion. The Committee noted that the developer assessed face validity systematically which met the testing requirement for a new measure and noted that the risk-adjustment model demonstrated adequate discrimination and calibration. However, the Committee expressed concerns that the variability of the underlying population could present a threat to validity. The Committee agreed that the measure is highly feasible to report, given that it is a claims-based measure. During the Use and Usability discussion, the Committee members raised concerns about the generalizability of the data and the impact that may

have on the usefulness of the measure. During the public comment period, the measure developer submitted a request for reconsideration of the Standing Committee's decision not to recommend this measure for endorsement due to concerns about the measure's validity. As the Standing Committee did not have quorum on its post comment call it was unable to discuss this request, an endorsement decision on the measure was deferred until the Spring 2019 cycle. During the spring 2019 review cycle, the Committee considered the request, but ultimately upheld its initial decision not to recommend the measure.

3445 All Cause Inpatient Admission Rate for Medicaid Beneficiaries with Complex Care Needs and High Costs (BCNs) (Mathematica Policy Research): Not Recommended

Description: All Cause inpatient admission rate for adult Medicaid beneficiaries who meet BCN population eligibility criteria. The measure is calculated as the number of inpatient admissions per 1,000 beneficiary months and is intended to be reported at the state level. For the purpose of this measure, the BCN population is defined as Medicaid beneficiaries who are age 18 to 64 during the lookback year (the 12 months prior to the measurement year) and the measurement year and have at least one inpatient admission and at least two chronic conditions, as defined by the Chronic Conditions Data Warehouse (CCW), during the lookback year. Beneficiaries dually enrolled in Medicaid and Medicare and beneficiaries who had fewer than 10 months of Medicaid eligibility in the lookback year are not included in the analytic sample because we did not have enough utilization data to include them in testing. We further limited the analytic file to beneficiaries that met the BCN definition criteria described above. **Measure Type**: Outcome; **Level of Analysis**: Population: Regional and State; **Setting of Care**: Inpatient/Hospital; **Data Source**: Claims

NQF 3445 measures an All Cause inpatient admissions rate for Medicaid beneficiaries between 18 and 64 years old who meet the criteria of presenting complex care needs and high costs. It was developed with the intention of pairing it with NQF 3443. The Committee agreed there was evidence that the measured entity could influence outcomes, citing evidence showing multiple interventions that could decrease inpatient utilization of complex patients. To demonstrate a performance gap, the developer cited both disparities in terms of race and ethnicity in performance for admission rates. The Committee also noted variation in performance across states. The developer provided conducted signal-to-noise (SNR) reliability testing using MAX data from 10 states. Committee members noted that the scores ranged from 0.95 to 0.99 and agreed that the measure was adequately reliable. The developer conducted convergent validity testing by examining the correlation between this measure and the HEDIS inpatient hospital utilization measure (IHU). However, the Committee raised concerns that the generalizability of the testing data threatened validity. The Committee agreed that the measure is highly feasible to report given that it is a claims-based measure. During the Use and Usability discussion, Committee members again raised concerns about the generalizability of the sample population to the larger Medicaid population and noted the potential for negative unintended consequences. During the public comment period, the measure developer submitted a request for reconsideration of the Standing Committee's decision not to recommend this measure for endorsement due to concerns about the measure's validity. As the Standing Committee did not have quorum on its post comment call it was unable to discuss this request, an endorsement decision on the measure was deferred until the spring

2019 cycle. During the spring 2019 review cycle, the Committee considered the request, but ultimately upheld its initial decision not to recommend the measure.

References

1 Joynt KE, Jha AK. Thirty-day readmissions — truth and consequences. *N Engl J Med*. 2012;366(15):1366-1369.

2 D McCarthy, Cohen A, Johnson MB. https://www.pcpcc.org/sites/default/files/1658_McCarthy_care_transitions_synthesis_v2.pdf. Last accessed July 2019.

3 Kripalani S, Chen G, Ciampa P, et al. A transition care coordinator model reduces hospital readmissions and costs. *Contemp Clin Trials*. 2019;81:55-61.

4 Boccuti C, Casillas G. Aiming for Fewer Hospital U-turns: The Medicare Hospital Readmission Reduction Program. *Henry J Kais Fam Found*. March 2017. https://www.kff.org/medicare/issue-brief/aiming-for-fewer-hospital-u-turns-the-medicare-hospital-readmission-reduction-program/. Last accessed July 2019.

5 Centers for Medicare & Medicaid Services (CMS). Merit-Based Incentive Payment System (MIPS): Scoring 101 Guide for Year 2 (2018). https://webcache.googleusercontent.com/search?q=cache:V-DffHKyfBgJ:https://qpp-cm-prod-content.s3.amazonaws.com/uploads/179/2018%2520MIPS%2520Scoring%2520Guide_Final.pdf+&cd=1&hl=en&ct=clnk&gl=us. Last accessed July 2019.

6 MedPAC. *Report to the Congress: Medicare and the Health Care Delivery System*. Washington, DC: Medicare Payment Advisory Commission; 2018.

7 Sabbatini AK, Wright B. Excluding observation stays from readmission rates — what quality measures are missing. *N Engl J Med*. 2018;378(22):2062-2065.

8 Dharmarajan K, Qin L, Bierlein M, et al. Outcomes after observation stays among older adult Medicare beneficiaries in the USA: retrospective cohort study. *BMJ*. 2017;357:j2616.

9 Bunis D. Need Rehab? Beware of Medicare's "Observation Status." AARP. http://www.aarp.org/health/medicare-insurance/info-2018/medicare-observation-rule-fd.html. Published February 2, 2018. Last accessed July 2019.

Appendix A: Details of Measure Evaluation

Rating Scale: H=High; **M**=Moderate; **L**=Low; **I**=Insufficient; **NA**=Not Applicable

Deferred Measure – for Clinician Group/Practice Specification Only

---

3495 Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate for the Merit-Based Incentive Payment System [Clinician Group/Practice Level of Analysis]

Submission | Specifications

Description: This measure is a re-specified version of the hospital-level measure, "Hospital-Wide All Cause, Unplanned Readmission Measure" (NQF 1789), which was developed for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare and are hospitalized in non-federal hospitals.

This re-specified measure attributes hospital-wide index admissions to up to three participating MIPS Eligible Clinicians or Eligible Clinician Groups ("providers"), rather than to hospitals. It assesses each provider's rate of 30-day readmission, which is defined as unplanned, All Cause readmission within 30 days of hospital discharge for any eligible condition.

The measure reports a single summary risk adjusted readmission rate (RARR), derived from the volume-weighted results of five different models, one for each of the following specialty cohorts based on groups of discharge condition categories or procedure categories: surgery/gynecology; general medicine; cardiorespiratory; cardiovascular; and neurology, each of which will be described in greater detail below.

Numerator Statement: The outcome for this measure is readmission within 30-days of a hospital discharge. We define readmission as an inpatient admission for any cause, except for certain planned readmissions, within 30 days from the date of discharge from an eligible index admission.

Denominator Statement: The measure includes admissions for Medicare beneficiaries who are 65 years and older and are discharged from any non-federal, acute care inpatient US hospitals (including territories) with Medicare part A enrollment for the 12 months prior to admission and part A enrollment for the 30 days after discharge. These are called 'index admissions'

Outcome attribution:

Each index admission is attributed to up to 3 eligible clinicians or eligible clinician groups.

1) One is the eligible clinician who filed a claim for the 'discharge procedure' (CPT code 99238 or 99239) for the patient; conceptually, this clinician is measured because, having billed for the discharge of the patient, they have some responsibility for the transition of the patient to non-acute settings.

2) Second is the eligible clinician who, during the inpatient stay, billed the most patient-facing charges; conceptually, this clinician has the most responsibility for the care of patients during their stay and may also be the Discharge Clinician.

3) Third is the eligible clinician that provides the plurality of outpatient primary care during the 12 months prior to the admission, as measured by plurality of primary care services; conceptually, a primary care provider may manage the transition from acute to non-acute care and participate in decisions to return to acute care.

Index admissions are attributed to a clinician by each of these rules; two or all three rules may attribute the index admission to the same clinician. Then, all admissions assigned to an eligible clinician are used

to construct a single measure score for that clinician, regardless of the reason the admission was attributed. The measure has also been tested for eligible clinician groups, implemented here by grouping eligible clinicians who use the same Taxpayer Identification Number (TIN).

Exclusions: From the cohort, we exclude admissions if:

1. The patient is discharged against medical advice (AMA)
2. The patient is discharged from a PPS-exempt cancer hospital
3. The patient is admitted primarily for the medical treatment of cancer
4. The patient is admitted primarily for the treatment of psychiatric disease
5. The patient is admitted primarily for "rehabilitation care; fitting of prostheses and adjustment devices" (CCS 254)
6. Admissions without 30 Days of Post-Discharge Enrollment are excluded
7. Admissions cannot be identified in IDR database
8. The admission cannot be attributed to an eligible clinician.

Further exclusion details can be found in S.9 Denominator Exclusion Details

Adjustment/Stratification: Statistical risk model

Level of Analysis: Clinician: Group/Practice

Setting of Care: Inpatient/Hospital

Type of Measure: Outcome

Data Source: Claims, Enrollment Data

Measure Steward: Centers for Medicare & Medicaid Services (CMS)

STANDING COMMITTEE MEETING [June 21, 2019]

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

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Y-14**; **N-2**; 1b. Performance Gap: **H-0**; **M-15**; **L-1**; **I-0**

Rationale:

- This is a re-specified version of the hospital-level measure, "Hospital-Wide All Cause, Unplanned Readmission Measure" (NQF 1789). NQF 1789 was developed for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare and are hospitalized in non-federal hospitals. This specified measure attributes admissions to up to three participating MIPS eligible clinicians.
- The Standing Committee reviewed the logic model presented by the developer demonstrating physician group interventions that can reduce the risk of unplanned hospital visits.
- The Standing Committee reviewed the range of performance for clinician groups from 13.1 in the first decile to 18.0 in the tenth decile.

---

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

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

2a. Reliability: **H-0**; **M-13**; **L-2**; **I-1**; 2b. Validity: **H-0**; **M-10**; **L-6**; **I-0**

Rationale:

- The Standing Committee reviewed the input provided by the NQF Scientific Methods Panel (SMP).
- The Standing Committee noted SMP concerns that social risk factors are excluded from the risk model given the effect size and the potential for negative consequences on access to care if this measure is not adequately risk adjusted. The Standing Committee agreed that the developer should examine other clinical variables that could underlie disparities such as frailty or functional status.
- The Standing Committee reviewed the measure score empirical validity testing and face validity testing submitted at the individual physician and physician group level. The Standing Committee noted that the evidence and validity testing must be evaluated separately for the two levels of analysis. The Committee agreed to vote individually on the two levels of analysis.
- The Standing Committee noted that eligible clinician groups' risk adjusted readmission rates go down with increasing Overall Hospital Quality Star Rating and with increasing quintile of the Star Rating readmission quality score.
- The Committee reviewed the Face Validity testing and results, the approach to risk adjustment and the conceptual model for socio-demographic risk adjustment.
- The Standing Committee had a mixed review of the face validity of the measure, specifically, the role physician groups have in improving this outcome.

---

3. Feasibility: H-6; M-9; L-1; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee agreed that the measure uses claims data that can be operationalized; however, the measure is not yet in use. There are no fees, licensing, or requirements to use the measure.

---

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-14; No Pass-2**; 4b. Usability: **H-1; M-11; L-4; I-0**

Rationale:

- The Standing Committee acknowledged that this measure is planned for use in the CMS MIPS program.
- The Standing Committee noted that this is a new measure and there is no information available on performance improvement. This measure is not currently used in a program, but a primary goal of the measure is to provide information necessary to implement focused quality improvement efforts. Once the measure is implemented, the developer plans to examine trends in improvements by comparing RSRR over time.

---

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Original Vote: N/A Final decision: Deferred**

6. Public and Member Comment

One comment was submitted during the post-evaluation comment period:

> The American Medication Association (AMA) appreciates the Standing Committee discussion and evaluation of this measure but continues to have significant concerns with the lack of adherence to the Consensus Development Process and whether the measure meets the NQF Measure Evaluation Criteria, particularly for evidence and scientific acceptability.
>
> The NQF has had a longstanding commitment to ensuring that the CDP and associated criteria are followed consistently and the process is conducted in a transparent manner. Unfortunately, we do not believe that it is demonstrated in this project and associated report. Specifically, the AMA is concerned with the limited number of members who were able to participate in the evaluation of this measure on the June 21 webinar; specifically, the roll call prior to discussion of this measure identified only 11 of the 21 members. Based on our review of the votes available for the individual clinician and group levels of analysis, an additional five members evaluated the measures against the criteria but were not present during the discussion of the measure on June 21. It is concerning to have just 50% of the committee participate in the public discussion of the measure and almost 25% of the remaining members participate in voting on a measure for which it is not clear they were able to fully evaluate, ask questions of the developer, and hear public comments. In addition, the draft report released for comment does not include the committee votes for feasibility, usability and use, and the recommendation for endorsement for the group level of analysis (see pages 13-14) but the narrative indicates that it is recommended for endorsement. Omissions like these lead us to question the integrity and consistency of the process and makes it extremely difficult for NQF members and the public to engage in the CDP in a meaningful way.

As mentioned in our comments submitted prior to the committee's evaluation, we believe that:

- Insufficient evidence was provided to support attribution of the measure to physicians or practices in the absence of some coordinated program or targeted intervention led by the health system or hospital;
- Assignment of responsibility of the reduction of readmissions to multiple physicians and practices in MIPS is not appropriate nor has the developer provided sufficient information to support the attribution of this measure to up to three physicians or practices;
- The measure score reliability results are too low when based on the minimum case number of 25 patients. Measures should meet minimum acceptable thresholds of 0.7 for reliability; and
- The conceptual basis used to explain which social risk factors were tested in Section 2b3.3a is inadequate and additional testing is needed to evaluate clinical factors in conjunction with social risk factors as well as the impact that the inclusion of these factors had on the absolute change of the rates.

As a result, the AMA is unable to support endorsement of the measure at this time and requests that NQF distribute the missing information in the report and the Committee reconsiders its recommendation for endorsement.

COMMITTEE RESPONSE:

The Standing Committee acknowledges the main concerns raised by the commenter related to: evidence to support attribution to physicians or practices, assignment of responsibility of readmissions to multiple physicians, reliability results at case numbers of less than 25 patients, and approach to testing of social risk factors.

The Committee agrees that improvement of readmissions requires shared accountability among all members of the care team but also struggled with the attribution to an individual clinician. Thus, the Committee agreed to spilt consideration of the measure based on individual clinician and clinician group level of analysis, noting different conceptual attribution issues and reliability performance at the two levels of analysis. The Committee generally agreed that individual physicians often rely on peers, office and hospital-based support staff for important care transition roles and may have limited direct influence on the social risk and community supports available to the patients they serve. However, physician group practices do have a role in improving readmission outcomes and are more likely able to provide the resources to improve care transitions.

On the issue of testing of social risk factors, the Committee reviewed testing information on the conceptual basis for social risk adjustment, the variables selected and tested, and the rationale for not including these variables in the final model risk model. The developed noted that the inclusion of the social risk factors in the final model did not meaningfully change risk model performance or measure score distribution. The Committee agrees that further future testing should be undertaken by the developer in accordance with the commenter recommendations.

Finally, the Committee reviewed the reliability testing data provided by the developer in the submission form. Committee members noted that the reliability methods and statistics submitted by the developer were considered by the NQF Scientific Methods Panel. Members of the Standing Committee agreed that the reliability statistics at lower case volume will demonstrate lower reliability, but the overall mean reliability statistics were acceptable. The Standing Committee requested the developer provide additional reliability statistics at lower case volume, both the range of performance and the mean values to improve transparency for all stakeholders. During the call, the Standing Committee generally agreed that the method and overall reliability performance was sufficient at the physician group level.

However, after reviewing the updated reliability testing information submitted via email after the call, and noting that incorrect reliability testing information was given verbally during the call which informed the Committee's discussion and voting, NQF, in consultation with the measure steward and developer and Standing Committee Co-chairs, elected to defer the final recommendation and send the measure back to the Committee for re-review in the Fall 2019 cycle.

NQF RESPONSE:

Thank you for your comments. NQF strives to achieve quorum at each step of the Consensus Development Process (CDP). Recognizing the burden on volunteer members of the CDP Committees with the increased activities in the bi-annual cycle of the CDP, NQF will be exploring process improvement opportunities to ensure future Committee calls do achieve quorum. With regards to voting, NQF followed our established procedure for cases when quorum is not achieved. Votes were not taken during the call where quorum was not reached. Following the call, the transcript and recording were provided to all Committee members along with a voting survey. Committee members who were not present are able to review the call materials and the transcript or recording prior to submitting their votes. Votes were accepted until quorum of the Committee was achieved.

There was an oversight by NQF staff in providing a full count of votes in the published version of the report. NQF corrected and reposted the report with the full information as soon as this comment was received. We thank the commenter for alerting us of this oversight.

MEASURE STEWARD/DEVELOPER RESPONSE:

We appreciate this summary of your earlier comments, which we address below.

We also agree with the conclusions outlined within NQF's final report, Improving Attribution Models (NQF, 2018), in that attribution models should reflect clinicians and providers with reasonable influence on the care and outcomes for patients in order to enforce accountability and facilitate quality improvement. During development, we solicited a wide variety of clinician, technical, and patient feedback through stakeholder engagement. The Technical Expert Panel, in particular, felt strongly that it was appropriate to attribute readmissions to multiple clinicians to encourage coordination and shared accountability. Additionally, the same panel identified the three clinicians attributed by this measure as being most accountable.

We agree that it is important that the final volume threshold correspond to adequate reliability. Constructing meaningful, reliable, valid provider quality measures is challenging and requires balancing competing factors and values. In the NQF Submission forms, we provide evidence that these measures do capture reliable and valid quality signals at the clinician and group level under the proposed attribution.

We tested for the effects of including two social risk factors within the model (dual eligibility status and low Agency for Healthcare Research and Quality SES) on final risk-adjusted rates for both clinicians and clinician groups. The correlation between the adjusted and unadjusted scores were 0.99, indicating extremely high agreement and that adding these social risk factors would have minimal impact on measure scores. Ongoing research aims to identify valid patient-level social risk factors and highlight disparities related to social risk. As additional variables become available, they will be considered for testing and inclusion within the measure. There are also alternative ways to adjust for social risk as part of measure program implementation, such as stratification or peer grouping, which CMS recently applied to the Hospital Readmission Reduction Program (HRRP).

Since the release of the Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors report in July 2017, NQF announced the launch of a new, three-year initiative to explore unresolved issues that surfaced in the 2015-2017 social risk factor trial.[a] The stated goal of the new Social Risk Trial is to "help inform a decision on whether to permanently change NQF's policy to allow social risk adjustment for outcome measures."[b] For risk-adjusted outcome measures, CMS first considers adjustment for clinical conditions and then examines additional risk imparted by social risk factors after the potential for greater disease burden is included in the risk model. We believe that this is consistent with NQF current guidance and is appropriate given the evidence cited in our submission that people who experience greater social risk are more likely to have more disease burden compared with those who do not; and that this is clearly not a signal of hospital quality. In addition, according to NQF guidance, developers should assess social risk factors for their contribution of unique variation in the outcome – that they are not redundant.[c] Therefore, if clinical risk factors explain all or most of the patient variation in the outcome, then NQF guidance does not support adding social risk factors that do not account for variation.

In addition to the correlation between adjusted and unadjusted scores, we also tested the change in risk-adjusted readmission rates. When incorporating the duel eligible risk factor, risk-adjusted readmission rates dropped an absolute value of 0.03% for clinicians and 0.02% for clinician groups. When incorporating low AHRQ SES, risk-adjusted readmission rates dropped an absolute value of -0.02% for clinicians and -0.01% for clinician groups.

NQF doesn't specify or require testing for impact on program inclusion, program benchmarking, or star rating systems. At this time, it is not known how CMS will use this measure in the MIPS program.

We agree with the importance of balancing these competing considerations. We are committed to constant refinement and improvement of risk adjustment models used in all measures. We will reevaluate this model and available risk factors on an ongoing basis, with the goal of

[a] National Quality Forum (NQF). NQF Statement on Board of Directors Decision Regarding Social Risk Trial, http://www.qualityforum.org/News_And_Resources/Press_Releases/2017/NQF_Statement_on_Board_of_Directors_Decision_Regarding_Social_Risk_Trial.aspx

[b] National Quality Forum (NQF). Social Risk Trial FAQ, June 28, 2018. http://www.qualityforum.org/ProjectMaterials.aspx?projectID=87820. Accessed September 9, 2019

[c] National Quality Forum (NQF). Risk adjustment for socioeconomic status or other sociodemographic factors: Technical report. 2014; http://www.qualityforum.org/Publications/2014/08/Risk_Adjustment_for_Socioeconomic_Status_or_Other_Sociodemographic_Factors.aspx. Accessed September 3, 2019.

producing the most accurate and fair risk adjustment models for assessing provider performance.

Updated Committee Recommendation: Deferred to Fall 2019 Cycle

Measure 3495 *Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate* was recommended at the Clinician: Group level of analysis. During the Committee's post-comment call on October 2, 2019, in response to questions concerning the testing and comments provided during the public comment period, the developer inadvertently provided incorrect information on the reliability results as the Committee was preparing to vote on whether to reconsider the measure. While the Committee did vote to uphold their recommendation for endorsement, the vote was extremely close (in order to formally reconsider a measure, more than 60% of the Committee must vote to do so; in this case, 57%, or 6 of the 14 Committee members present, voted to reconsider). After the call, the developer provided the correct reliability results in writing (which had also been provided in their original submission). In consultation with the developers and the Committee co-chairs, NQF staff decided to return measure 3495 to the Committee for re-evaluation at the Clinician Group level of analysis, during the Fall 2019 cycle, due to concerns that the incorrect information provided may have influenced the Committee's vote.

7. Consensus Standards Approval Committee (CSAC) Endorsement Decision (October 21, 2019): N/A

Though not required to vote on the decision to return the measure for evaluation, the CSAC agreed with the rationale to send 3495 back to the Standing Committee during the Fall 2019 cycle at the Clinician Group/practice level of analysis only.

8. Appeals:

N/A

3495 Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate for the Merit-Based Incentive Payment System [Individual Clinician Level of Analysis]

Submission | Specifications

Description: This measure is a re-specified version of the hospital-level measure, "Hospital-Wide All Cause, Unplanned Readmission Measure" (NQF 1789), which was developed for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare and are hospitalized in non-federal hospitals.

This re-specified measure attributes hospital-wide index admissions to up to three participating MIPS Eligible Clinicians or Eligible Clinician Groups ("providers"), rather than to hospitals. It assesses each provider's rate of 30-day readmission, which is defined as unplanned, All Cause readmission within 30 days of hospital discharge for any eligible condition.

The measure reports a single summary risk adjusted readmission rate (RARR), derived from the volume-weighted results of five different models, one for each of the following specialty cohorts based on groups of discharge condition categories or procedure categories: surgery/gynecology; general medicine; cardiorespiratory; cardiovascular; and neurology, each of which will be described in greater detail below.

Numerator Statement: The outcome for this measure is readmission within 30-days of a hospital discharge. We define readmission as an inpatient admission for any cause, except for certain planned readmissions, within 30 days from the date of discharge from an eligible index admission.

Denominator Statement: The measure includes admissions for Medicare beneficiaries who are 65 years and older and are discharged from any non-federal, acute care inpatient US hospitals (including territories) with Medicare part A enrollment for the 12 months prior to admission and part A enrollment for the 30 days after discharge. These are called 'index admissions'

Outcome attribution:

Each index admission is attributed to up to 3 eligible clinicians or eligible clinician groups.

1) One is the eligible clinician who filed a claim for the 'discharge procedure' (CPT code 99238 or 99239) for the patient; conceptually, this clinician is measured because, having billed for the discharge of the patient, they have some responsibility for the transition of the patient to non-acute settings.

2) Second is the eligible clinician who, during the inpatient stay, billed the most patient-facing charges; conceptually, this clinician has the most responsibility for the care of patients during their stay and may also be the Discharge Clinician.

3) Third is the eligible clinician that provides the plurality of outpatient primary care during the 12 months prior to the admission, as measured by plurality of primary care services; conceptually, a primary care provider may manage the transition from acute to non-acute care and participate in decisions to return to acute care.

Index admissions are attributed to a clinician by each of these rules; two or all three rules may attribute the index admission to the same clinician. Then, all admissions assigned to an eligible clinician are used to construct a single measure score for that clinician, regardless of the reason the admission was attributed. The measure has also been tested for eligible clinician groups, implemented here by grouping eligible clinicians who use the same Taxpayer Identification Number (TIN).

Exclusions: From the cohort, we exclude admissions if:

1. The patient is discharged against medical advice (AMA)
2. The patient is discharged from a PPS-exempt cancer hospital
3. The patient is admitted primarily for the medical treatment of cancer
4. The patient is admitted primarily for the treatment of psychiatric disease
5. The patient is admitted primarily for "rehabilitation care; fitting of prostheses and adjustment devices" (CCS 254)
6. Admissions without 30 Days of Post-Discharge Enrollment are excluded
7. Admissions cannot be identified in IDR database
8. The admission cannot be attributed to an eligible clinician.

Further exclusion details can be found in S.9 Denominator Exclusion Details

Adjustment/Stratification: Statistical risk model

Level of Analysis: Clinician: Individual

Setting of Care: Inpatient/Hospital

Type of Measure: Outcome

Data Source: Claims, Enrollment Data

Measure Steward: Centers for Medicare & Medicaid Services (CMS)

STANDING COMMITTEE MEETING [June 21, 2019]

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

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Y-11**; **N-5**; 1b. Performance Gap: **H-0**; **M-11**; **L-4**; **I-1**;

Rationale:

- The Standing Committee reviewed the distribution of risk adjusted readmission ratios (RARRs) for eligible clinician's ranges from a mean of 11.4 in the first decile to 19.5 in the tenth decile.

---

2. Scientific Acceptability of Measure Properties: The measure does not meet the Scientific Acceptability criteria

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

2a. Reliability: **H-0**; **M-8**; **L-5**; **I-3**; 2b. Validity: **H-0**; **M-5**; **L-10**; **I-1**

Rationale:

- Similar to the review of the physician-group specification of this measure, the Standing Committee reviewed the input provided by the NQF Scientific Methods Panel (SMP).
- The Standing Committee reviewed the reliability assessment of the SMP subgroup members.
- Similar to the Validity summary provided above the Standing Committee reviewed the validity testing provided by the developer, the risk adjustment methodology, and the calibration statistics.
- The Standing Committee raised several concerns about the individual clinician specifications and testing. Specifically, the Committee was concerned about the attribution approach. The face validity of holding an individual clinician accountable for All Cause hospital readmission raised concern with the Standing Committee. The Committee was not clear how an individual physician could directly influence these outcomes without collaboration of other physicians and hospital

partnerships.

3. Feasibility: H-X; M-X; L-X; I-X: The Standing Committee did not vote on this criteria since the measure did not pass Scientific Acceptability

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee generally agreed that the measure uses claims data that can be operationalized; however, the measure is not yet in use. There are no fees, licensing, or requirements to use the measure.

4. Usability and Use: The Standing Committee did not vote on this criteria since the measure did not pass Scientific Acceptability

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-X; No Pass-X**; 4b. Usability: **H-X; M-X; L-X; I-X**

Rationale:

- The Standing Committee acknowledged that this measure is planned for use in the CMS MIPS program.
- The Standing Committee had concerns about the usability of a clinician-level readmission measure to information necessary to implement focused quality improvement efforts. The Standing Committee noted that the benefits of the performance measure may not outweigh evidence of unintended negative consequences to individuals or populations given the locus of control of an individual clinician to the outcome.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: Yes-X; **No-X**

Rationale

- The Standing Committee did not vote on this criteria since the measure did not pass Scientific Acceptability.

6. Public and Member Comment

NQF received one comment on this measure during the post-measure evaluation meeting commenting period. The commenter noted inconsistences in the level of committee member participation across several measure evaluation meetings and called attention to the need for consistent quorum during evaluation meetings. The commenter also identified omissions in the Spring 2019 draft report for comment. The commenter expressed that the measure should not be endorsed due to concerns over insufficient evidence to support attribution of the measure to physicians or practices in the absence of targeted intervention or coordinated programs led by the health system or hospital. Additional concerns were expressed regarding the measure's score reliability results and its attribution to multiple physicians

and practices. Lastly, the commenter believed the measure did not adequately qualify, test or adjust for social risk factors and were concerned about what was included or excluded in the measure.

NQF Response:
Thank you for your comments. NQF strives to achieve quorum at each step of the Consensus Development Process (CDP). Recognizing the burden on volunteer members of the CDP Committees with the increased activities in the bi-annual cycle of CDP activities, NQF will be exploring process improvement opportunities to ensure future Committee calls do achieve quorum. With regards to voting, NQF followed our established procedure for cases when quorum is not achieved. Votes were not taken during the during the call. Following the call, the transcript and recording was provided to all Committee members along with a voting survey. Committee members who were not present are able to review the call materials and the transcript or recording prior to submitting their votes. Votes were accepted until quorum of the Committee was achieved.

There was an oversight by NQF staff in providing a full count of votes in the published version of the report. NQF corrected and reposted the report with the full information as soon as this comment was received. We thank the Commenter for alerting us of this oversight.

Developer Response:
Yale CORE agrees with the conclusions outlined within NQF's final report, Improving Attribution Models (NQF, 2018), in that attribution models should reflect clinicians and providers with reasonable influence on the care and outcomes for patients in order to enforce accountability and facilitate quality improvement. During development, we solicited a wide variety of clinician, technical, and patient feedback through stakeholder engagement. The Technical Expert Panel, in particular, felt strongly that it was appropriate to attribute readmissions to multiple clinicians to encourage coordination and shared accountability. Additionally, the same panel identified the three clinicians attributed by this measure as being most accountable.

Yale CORE agrees that it is important that the final volume threshold correspond to adequate reliability. Constructing meaningful, reliable, valid provider quality measures is challenging and requires balancing competing factors and values. In the NQF Submission forms, we provide evidence that these measures do capture reliable and valid quality signals at the clinician and group level under the proposed attribution.

Yale CORE tested for the effects of including two social risk factors within the model (dual eligibility status and low Agency for Healthcare Research and Quality SES) on final risk-adjusted rates for both clinicians and clinician groups. The correlation between the adjusted and unadjusted scores were 0.99, indicating extremely high agreement and that adding these social risk factors would have minimal impact on measure scores. Ongoing research aims to identify valid patient-level social risk factors and highlight disparities related to social risk. As additional variables become available, they will be considered for testing and inclusion within the measure. There are also alternative ways to adjust for social risk as part of measure program implementation, such as stratification or peer grouping, which CMS recently applied to the Hospital Readmission Reduction Program (HRRP).

Since the release of the Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors report in July 2017, NQF announced the launch of a new, three-year initiative to explore unresolved issues that surfaced in the 2015-2017 social risk factor trial. The stated goal of the new Social Risk Trial is to "help inform a decision on whether to permanently change NQF's policy to allow social risk adjustment for outcome measures." For risk-adjusted outcome measures, CMS first considers adjustment for clinical conditions and then examines additional risk imparted by social risk factors after the potential for greater disease burden is included in the risk model. We believe that this is consistent with NQF current guidance and is appropriate given the evidence cited in our submission that people who experience greater social risk are more likely to have more disease burden compared with those who do not; and that this is clearly not a signal of hospital quality. In addition, according to NQF guidance, developers should assess social risk factors for their contribution of unique variation in the outcome – that they are not redundant. Therefore, if clinical risk factors explain all or most of the patient variation in the outcome, then NQF guidance does not support adding social risk factors that do not account for variation.

In addition to the correlation between adjusted and unadjusted scores, Yale CORE also tested the change in risk-adjusted readmission rates. When incorporating the duel eligible risk factor, risk-adjusted readmission rates dropped an absolute value of 0.03% for clinicians and 0.02% for clinician groups. When incorporating low AHRQ SES, risk-adjusted readmission rates dropped an absolute value of -0.02% for clinicians and -0.01% for clinician groups.

NQF doesn't specify or require testing for impact on program inclusion, program benchmarking, or star rating systems. At this time, it is not known how CMS will use this measure in the MIPS program.

Yale CORE agrees with the importance of balancing these competing considerations. We are committed to constant refinement and improvement of risk adjustment models used in all measures. We will reevaluate this model and available risk factors on an ongoing basis, with the goal of producing the most accurate and fair risk adjustment models for assessing provider performance.

7. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Yes-12; No-0 (October 21, 2019): Decision: Not recommended for endorsement. The CSAC voted unanimously to uphold the Committee's recommendation not to endorse 3495 at the individual clinician level of analysis and noted no concerns.

---

8. Appeals

N/A

3443 All Cause Emergency Department Utilization Rate for Medicaid Beneficiaries with Complex Care Needs and High Costs (BCNs)

Submission

Description: All Cause emergency department (ED) utilization rate for adult Medicaid beneficiaries who meet BCN population eligibility criteria. The measure is calculated as the number of ED visits per 1,000 beneficiary months and is intended to be reported at the state level.

For the purpose of this measure, the BCN population is defined as Medicaid beneficiaries who are age 18 to 64 during the lookback year (the 12 months prior to the measurement year) and the measurement year and have at least one inpatient admission and at least two chronic conditions, as defined by the Chronic Conditions Data Warehouse (CCW), during the lookback year. Beneficiaries dually enrolled in Medicaid and Medicare and beneficiaries who had fewer than 10 months of Medicaid eligibility in the lookback year are not included in the analytic sample because we did not have enough utilization data to include them in testing. We further limited the analytic file to beneficiaries that met the BCN definition criteria described above.

Numerator Statement: The number of ED visits in the measurement year among adult Medicaid beneficiaries who meet BCN population eligibility criteria.

Denominator Statement: Number of Medicaid-eligible months ("beneficiary months") among adult Medicaid beneficiaries who meet BCN population eligibility criteria.

Exclusions: N/A

Adjustment/Stratification: Statistical risk model

Level of Analysis: Population : Regional and State

Setting of Care: Emergency Department and Services

Type of Measure: Outcome

Data Source: Claims

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

STANDING COMMITTEE MEETING [2/7/2019]

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

(1a. Evidence: 1b. Performance Gap)

1a. Evidence: **Y-17; N-0** 1b. Performance Gap: **H-4; M-13; L-0; I-0**

Rationale:

- This measure of emergency department utilization for Medicaid beneficiaries with complex care needs and high costs (BCNs) assesses a heterogeneous population with disproportionately high use of inpatient and ED use.
- The developer noted that improvement on this outcome may involve strengthening beneficiaries' relationships with health care providers in the community, improved care coordination, and chronic disease management.
- The developer demonstrates an adjusted performance range of 109.5 admissions per 1,000 beneficiary months to 322.0 admissions per 1,000 beneficiary months.
- The Committee agreed there was evidence that the measured entity could influence the outcome. Specifically, the developer cited several studies demonstrating that emergency

department visits in complex patients could be reduced through improved care management and agreed there was variation in performance. The Committee also noted that emergency department use at the population or plan level is directly related to the inability to access care in the community.
- The Committee generally agreed that there was a performance gap in the focus area of this measure.

---

2. Scientific Acceptability of Measure Properties: <u>The measure does not meet the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-3; M-11; L-1; I-0** 2b. Validity: **H-0; M-3; L-8; I-4**

<u>Rationale</u>:

- The developer conducted signal-to-noise (SNR) reliability testing for this measure using MAX data from 10 states. Average signal-to-noise reliability estimate was 0.92 (ranging between 0.59 to 0.99 across the ten states in the sample).
- Committee members raise concerns about the measure's generalizability to all 50 states, given the representativeness of the data used in testing.
- The Committee raised several concerns under the validity sub-criterion. The Committee noted that the developer tested the validity of the measure using a face validity test. The Committee was concerned that only 11 out of the 17 TEP members responded to whether the measure was a good indicator of quality.
- The risk-adjustment approach was developed using data from 10 states. The risk-adjustment model included 69 risk factors. While the measure demonstrated adequate discrimination and calibration of the risk adjustment model, the Committee expressed concerns that the variability of the underlying patient population could present a threat to validity.
- The developer noted that they included all predictors that were theoretically associated with the measure, including those that were not statistically significant or "protective" in nature. The developer stated that in general, the risk factors associated with a lower adjusted risk of ED utilization reflected more serious conditions (e.g. colorectal cancer). This lower risk likely reflects higher substitution away from ED care towards inpatient care. Because BCN-1 will ultimately be paired with a measure of inpatient care, the developer believed it was important to include the "protective" risk factors in the BCN-1 risk adjustment model.
- The data set used to develop the measure is not necessarily representative of the country or the population that the measure is intended to be applied. Of the 50 states, 34 were excluded due to data issues.
- Further, the Committee discussed that there is significant variation in the eligible population for this measure due to the differences in the Medicaid populations between states. Thus, applying this measure to the heterogeneous Medicaid populations across states makes differences in measure performance across states difficult to interpret. Are the differences due to actual health system performance differences or are the differences due to underlying differences in the Medicaid populations? The inability for the developer to distinguish this brought into question the validity of the measure as currently constructed.
- Committee members did agree there are real differences in both performance and quality between states, but ultimately believed that the threats to validity were too strong and the measure did not pass this criterion.

3. Feasibility: H-X; M-X; L-X; I-X: The Standing Committee did not vote on this criteria since the measure did not pass Scientific Acceptability

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

Rationale:

- The Committee agreed the measure is feasible to report given that it is a claims-based measure. The Committee did note concerns with Medicaid churn since the measure is constructed to include patients who had coverage for ten months, and that may exclude a large number of people in some states.

4. Use and Usability: The Standing Committee did not vote on this criteria since the measure did not pass Scientific Acceptability

(4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Y-X; N-X** 4b. Usability: **H-X; M-X; L-X; I-X**

Rationale:

- During the Use and Usability discussion, the Committee members raised concerns about the generalizability of the data and the impact that may have on the usefulness of the measure.

5. Related and Competing Measures

- No related or competing measures noted.

6. Standing Committee Recommendation for Endorsement: **Y-X**; **N-X**

Rationale

- The Standing Committee does not recommend this measure for endorsement due to concerns about the measure's validity. Because quorum was not reached during the meeting, each criterion was discussed and voting was completed later. Validity is a must-pass criterion; therefore, subsequent votes for remaining criteria are not captured.

7. Public and Member Comment

- The steward and developer of this measure submitted a request for reconsideration based on an inappropriate application of the validity subcriterion.
- The measure steward and developer responded to the Committee's concerns about the measures' validity. Specifically, they commented on the Committee's concern about differences in Medicaid populations across states, whether the measure was tested with a representative data sample, and the data quality.
- The measure steward and developer provided several clarifications regarding the differences in Medicaid populations across states. They recognized that state Medicaid programs vary substantially both in the covered populations and the quality of data reported to CMS. However, these variations are due to the design of Medicaid of federal-state partnership, and the

developer raised concerns that the Committee's emphasis on state variation in Medicaid program created an unrealistic standard for validity.

- The steward and developer noted that they believe the measure was tested using a robust data sample for assessing measure performance. They noted the states providing data varied in location, geography, size, and delivery system (fee-for-service or managed care) while still providing high-quality data. Additionally, they commented that the differences across Medicaid programs due to eligibility policies, mix of delivery models, payment rates, and other features, make it challenging for any sample to be representative of all 50 states. The steward and developer commented that the goal of measure testing is to select a diverse group of states that have high-quality data and whose populations capture, for the key variables in question, the majority of the variation that also occurs within other states. They also clarified that the measure specifications were designed to maximize the likelihood that states could define the denominator population consistently.
- Finally, the steward and developer responded to the Committee's concerns about data quality. The developer worked to evaluate the quality of relevant data in all the states and selected those states whose data met our quality standards. Specifically, the states chosen for testing had indicators that aligned with national inpatient and emergency utilization benchmarks and did not have data anomalies that would raise analytic problems (such as high levels of missing data). Further, NQF has endorsed measures in the past that were tested with Medicaid data from the same data source. The measure steward and developer noted that they do not believe that state variation in data quality should be a key factor in determining suitability of Medicaid measures for endorsement as long as the data used for development is of sufficient quality. Additionally, they noted that even perfect data from all states will not change the fact that state Medicaid programs have differences in design and operational features.
- The Standing Committee did not have quorum on the post comment call and was unable to discuss the request for reconsideration.
- To allow the Standing Committee time to consider the request for reconsideration, an endorsement decision on the measure has been deferred until the Spring 2019 review cycle.
- During the Spring 2019 measure evaluation webinars, which were held on June 20 and 21, 2019, quorum was not achieved. As a result, the Committee voted on whether or not to reconsider the measures via an online SurveyMonkey. After a review of the submitted materials, the Committee generally agreed that the additional information did not adequately address their concerns. Committee members noted challenges with the data variability and a more standardized data collection approach was required for the measures to be ready for endorsement. The Committee voted not to reconsider (Y-4; N-14).

---

8. Consensus Standards Approval Committee (CSAC) Vote: Y-7; N-1 (November 12, 2019):

Decision: Not approved for endorsement.

3445 All Cause Inpatient Admission Rate for Medicaid Beneficiaries with Complex Care Needs and High Costs (BCNs)

Submission

Description: All Cause inpatient admission rate for adult Medicaid beneficiaries who meet BCN population eligibility criteria. The measure is calculated as the number of inpatient admissions per 1,000 beneficiary months and is intended to be reported at the state level.

For the purpose of this measure, the BCN population is defined as Medicaid beneficiaries who are age 18 to 64 during the lookback year (the 12 months prior to the measurement year) and the measurement year and have at least one inpatient admission and at least two chronic conditions, as defined by the Chronic Conditions Data Warehouse (CCW), during the lookback year. Beneficiaries dually enrolled in Medicaid and Medicare and beneficiaries who had fewer than 10 months of Medicaid eligibility in the lookback year are not included in the analytic sample because we did not have enough utilization data to include them in testing. We further limited the analytic file to beneficiaries that met the BCN definition criteria described above.

Numerator Statement: The sum of unique inpatient admissions and observation stays in the measurement year among adult Medicaid beneficiaries who meet BCN population eligibility criteria.

Denominator Statement: Number of Medicaid-eligible months ("beneficiary months") among adult Medicaid beneficiaries who meet BCN population eligibility criteria.

Exclusions: Not applicable

Adjustment/Stratification: Statistical risk model

Level of Analysis: Population: Regional and State

Setting of Care: Inpatient/Hospital

Type of Measure: Outcome

Data Source: Claims

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

STANDING COMMITTEE MEETING [2/7/2019]

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

(1a. Evidence: 1b. Performance Gap)

1a. Evidence: **Y-17; N-0** 1b. Performance Gap: **H-4; M-13; L-0; I-0**

Rationale:

- The Committee agreed that the issues raised for #3443 are similar to the issues for #3445.
- The Committee agreed there was evidence the measure entity could influence the outcome, citing evidence showing multiple interventions that could decrease inpatient utilization of complex patients with appropriate managed care.
- To support the evidence of a performance gap, the developers cited both disparities in terms of race and ethnicity in performance for admission rates.
- The Committee also noted variation in performance across states.

2. Scientific Acceptability of Measure Properties: <u>The measure does not meet the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-3; M-12; L-0; I-0** 2b. Validity: **H-0; M-6; L-7; I-2**

<u>Rationale</u>:

- The developer conducted signal-to-noise (SNR) reliability testing using MAX data from 10 states. Committee members noted the scores ranged from 0.95 to 0.99 and agreed that the measure demonstrated adequate reliability testing.
- The developer conducted convergent validity testing by examining the correlation between this measure and the HEDIS inpatient hospital utilization measure (IHU).
- The Committee raised a number of points about the validity of this measure. The data set used to develop the measure is not necessarily representative of the country or the population that the measure is intended to be applied. Of the 50 states, 34 were excluded due to data issues.
- Further, the Committee discussed the significant variation in the eligible population for this measure between states – due to the differences in the Medicaid populations between states. Thus, applying this measure to the heterogeneous Medicaid populations across states makes differences in measure performance across states difficult to interpret. Are the differences due to actual health system performance differences or are the differences due to underlying differences in the Medicaid populations? The inability for the developer to distinguish this brought into question the validity of the measure as currently constructed.
- Committee members did agree there are real differences in both performance and quality between states, but ultimately believed that the threats to validity were too strong and the measure did not pass this criterion.

3. Feasibility: H-X; M-X; L-X; I-X: <u>The Standing Committee did not vote on this criteria since the measure did not pass Scientific Acceptability</u>

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

<u>Rationale</u>:

- The Committee agreed the measure is highly feasible to report, given that it is a claims-based measure.

4. Use and Usability: <u>The Standing Committee did not vote on this criteria since the measure did not pass Scientific Acceptability</u>

(4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Y-X; N-X** 4b. Usability: **H-X; M-X; L-X; I-X**

<u>Rationale</u>:

- During the Use and Usability discussion, the Committee members again raised concerns about the generalizability of the data and noted the potential for negative unintended consequences.

5. Related and Competing Measures

- No related or competing measures noted.

---

6. Standing Committee Recommendation for Endorsement: **Y-X**; **N-X**

Rationale

- The Standing Committee does not recommend this measure for endorsement due to concerns about the measure's validity. Because quorum was not reached during the meeting, each criterion was discussed, and voting was completed later. Validity is a must-pass criterion; therefore, subsequent votes for remaining criteria are not captured.

---

7. Public and Member Comment

- The steward and developer of this measure submitted a request for reconsideration based on an inappropriate application of the validity subcriterion.
- The measure steward and developer responded to the Committee's concerns about the measures' validity. Specifically, they commented on the Committee's concern about differences in Medicaid populations across states, whether the measure was tested with a representative data sample, and the data quality.
- The measure steward and developer provided several clarifications regarding the differences in Medicaid populations across states. They recognized that state Medicaid programs vary substantially both in the covered populations and the quality of data reported to CMS. However, these variations are due to the design of Medicaid of federal-state partnership, and the developer raised concerns that the Committee's emphasis on state variation in Medicaid program created an unrealistic standard for validity.
- The steward and developer noted that they believe the measure was tested using a robust data sample for assessing measure performance. They noted the states providing data varied in location, geography, size, and delivery system (fee-for-service or managed care) while still providing high-quality data. Additionally, they commented that the differences across Medicaid programs due to eligibility policies, mix of delivery models, payment rates, and other features, make it challenging for any sample to be representative of all 50 states. The steward and developer commented that the goal of measure testing is to select a diverse group of states that have high-quality data and whose populations capture, for the key variables in question, the majority of the variation that also occurs within other states. They also clarified that the measure specifications were designed to maximize the likelihood that states could define the denominator population consistently.
- Finally, the steward and developer responded to the Committee's concerns about data quality. The developer worked to evaluate the quality of relevant data in all the states and selected those states whose data met our quality standards. Specifically, the states chosen for testing had indicators that aligned with national inpatient and emergency utilization benchmarks and did not have data anomalies that would raise analytic problems (such as high levels of missing data). Further, NQF has endorsed measures in the past that were tested with Medicaid data from the same data source. The measure steward and developer noted that they do not believe that state variation in data quality should be a key factor in determining suitability of Medicaid

measures for endorsement as long as the data used for development is of sufficient quality. Additionally, they noted that even perfect data from all states will not change the fact that state Medicaid programs have differences in design and operational features.

- The Standing Committee did not have quorum on the post comment call and was unable to discuss the request for reconsideration.
- To allow the Standing Committee time to consider the request for reconsideration, an endorsement decision on the measure has been deferred until the Spring 2019 review cycle.
- During the Spring 2019 measure evaluation webinars, which were held on June 20 an 21, 2019, quorum was not achieved. As a result, the Committee voted on whether or not to reconsider the measures via an online SurveyMonkey. After a review of the submitted materials, the Committee generally agreed that the additional information did not adequately address their concerns. Committee members noted challenges with the data variability and a more standardized data collection approach was required for the measures to be ready for endorsement. The Committee voted not to reconsider (Y-4; N-14).

---

8. Consensus Standards Approval Committee (CSAC) Vote: Y-7; N-1 (November 12, 2019):

Decision: Not approved for endorsement.

Withdrawn Measure

2539 Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy

Submission

Description: Facility-level risk-standardized rate of acute, unplanned hospital visits within 7 days of a colonoscopy procedure performed at a hospital outpatient department (HOPD) or ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older. An unplanned hospital visit is defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission. The measure is calculated separately for ASCs, and HOPDs.

Numerator Statement: Unplanned hospital visits within 7 days of a qualifying colonoscopy.

Denominator Statement: Colonoscopies performed at hospital outpatient departments (HOPDs) and ambulatory surgical centers (ASCs) for Medicare FFS patients aged 65 years and older.

Exclusions: S.4. Numerator Statement: Unplanned hospital visits within 7 days of a qualifying colonoscopy.

S.6. Denominator Statement: Colonoscopies performed at hospital outpatient departments (HOPDs) and ambulatory surgical centers (ASCs) for Medicare FFS patients aged 65 years and older.

S.8. Denominator Exclusions: We established the following exclusion criteria after reviewing the literature, examining existing measures, discussing alternatives with the working group and technical expert panel (TEP) members, and reviewing feedback from the national dry run held in July 2015. The goal was to be as inclusive as possible; we excluded only those high-risk procedures and patient groups for which risk adjustment would not be adequate or for which hospital visits were not typically a quality signal. The exclusions, based on clinical rationales, prevent unfair distortion of performance results.

1) Colonoscopies for patients who lack continuous enrollment in Medicare FFS Parts A and B in the 7 days after the procedure.

Rationale: We exclude these patients to ensure full data availability for outcome assessment.

2) Colonoscopies that occur concurrently with high-risk upper gastrointestinal (GI) endoscopy procedures.

Rationale: Patients undergoing concurrent high-risk upper GI endoscopy procedures, such as upper GI endoscopy procedures for the control of bleeding or treatment of esophageal varices, and have a higher risk profile than typical colonoscopy patients. Therefore, these patients have a disproportionally higher risk for the outcome.

3) Colonoscopies for patients with a history of inflammatory bowel disease (IBD) or diagnosis of IBD at time of index colonoscopy or on the subsequent hospital visit outcome claim.

Rationale: We exclude these patients because:

-IBD is a chronic condition; patients with IBD undergo colonoscopy both for surveillance due to increased cancer risk and for evaluation of acute symptoms. IBD is likely to be coded as the primary diagnosis prompting the procedure irrespective of whether the patients are undergoing a screening procedure or a diagnostic procedure in the setting of an acute exacerbation of IBD. Therefore, we may not be able to adequately risk adjust for these patients, as we cannot identify relatively well versus acutely unwell patients among visits coded as IBD.

-Our aim is to capture hospital visits which reflect the quality of care. Admissions for acutely ill IBD patients who are evaluated with an outpatient colonoscopy and are subsequently admitted for medical treatment of an IBD flare do not reflect the quality of the colonoscopy. In our 2010 Medicare 20% FFS

Full Development Sample (see the 2014 Facility 7-day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy Measure Technical Report posted at https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1228775197506) for full description of the dataset), more than one-third of IBD patients admitted to the hospital with colonoscopy had a discharge diagnosis of IBD, indicating their admission was for medical treatment of their IBD. We therefore excluded this group so that providers who treat a disproportionate number of IBD patients will not be disadvantaged in the measure.

-A post-index diagnosis of IBD, which represents a very small fraction of cases (less than 0.5% of the cohort) in the measure population, indicates that the condition was likely present at the time of the index colonoscopy but not coded.

4) Colonoscopies for patients with a history of diverticulitis or diagnosis of diverticulitis at time of index colonoscopy or on the subsequent hospital visit outcome claim.

Rationale: We exclude these patients because:

-It is unclear what the health status is of patients coded with a history or current diagnosis of diverticulitis, making it difficult to fully risk adjust for patients' health. Colonoscopies performed on patients with a history or current diagnosis of diverticulitis are likely to be coded as diverticulitis as the primary diagnosis irrespective of whether the patients are undergoing a screening procedure or a diagnostic procedure (i.e., are acutely unwell with active disease). Furthermore, the codes for diverticulitis and diverticulosis may not be consistently used; patients with diverticulosis may be erroneously coded as diverticulitis. Therefore, we may not be able to adequately risk adjust as we cannot identify relatively well versus acutely unwell patients among visits coded as diverticulitis.

-Admissions for acutely ill patients with a history or current diagnosis of diverticulitis who are evaluated with an outpatient colonoscopy and are subsequently admitted for medical treatment of do not reflect the quality of the colonoscopy. In our 2010 Medicare 20% FFS Full Development Sample (see the Facility 7-day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy Measure Technical Report posted on the web page provided in data field S.1) more than one-quarter of patients with a history or current diagnosis of diverticulitis admitted to the hospital post colonoscopy had a discharge diagnosis of diverticulitis, indicating they were admitted for medical treatment of the condition. These admissions are likely unrelated to the quality of the colonoscopy. We therefore excluded this group so that providers who treat a disproportionate number of diverticulitis patients will not be disadvantaged in the measure.

-A post-index diagnosis of diverticulitis, which represents a very small fraction of cases (less than 0.5% of the cohort) in the measure population, indicates that the condition was likely present at the time of the index colonoscopy but not coded.

5) Colonoscopies that occur on the same hospital claim as an ED visit, unless the ED visit has a diagnosis indicative of a complication of care (applies to colonoscopies at HOPDs only).

Rationale: We exclude these patients because:

-The sequence of events in these cases is not clear. It is not possible to use claims data to determine whether the colonoscopy was the cause of, subsequent to, or during the ED visit.

6) Colonoscopies that occur on the same day and at the same hospital as an ED visit that is billed on a separate claim than the index colonoscopy, unless the ED visit has a diagnosis indicative of a complication of care (applies to colonoscopies at HOPDs only).

Rationale: We exclude these patients because:

-It is unclear whether the same-day ED visit occurred before or after the colonoscopy. However, for ED visits billed on the same day but at a different facility, it is unlikely that a patient would experience an ED

visit for an acute diagnosis at one facility and then travel to another facility for a routine colonoscopy on the same day. Therefore, these colonoscopies are not excluded because they likely represent a routine procedure followed by a complication of care.

7) Colonoscopies that occur on the same hospital outpatient claim as an observation stay (applies to colonoscopies at HOPDs only).

Rationale: We exclude these patients because:

-The sequence of events in these cases is not clear. It is not possible to use claims data to determine whether the colonoscopy was the cause of, subsequent to, or during the observation stay.

8) Colonoscopies followed by a subsequent outpatient colonoscopy procedure within 7 days.

Rationale: We exclude these patients because:

-The two colonoscopies are considered part of a single episode of care, for which the subsequent colonoscopy is considered the index procedure.

Adjustment/Stratification: Statistical risk model

Level of Analysis: Facility

Setting of Care: Outpatient Services

Type of Measure: Outcome

Data Source: Claims, Other

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING [June 20, 2019]

1. Importance to Measure and Report: <u>The Standing Committee did not vote as the measure was withdrawn</u>

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Y-X; N-X**; 1b. Performance Gap: **H-X**; **M-X**; **L-X**; **I-X;**

<u>Rationale</u>:

- The Standing Committee reviewed this measure of hospital visits within 7 days of a colonoscopy procedure performed at a hospital outpatient department (HOPD) or ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older captures adverse patient outcomes associated with HOPD and ASC care and an important area for quality improvement.
- The Standing Committee agreed that the measure provides a logic model demonstrating provider-level and facility-level interventions that can reduce the risk of unplanned hospital visits.
- These provider-level factors include protocol for patient's colonoscopy prep, and technical quality of the procedure. The facility-level factors include anesthesia, discharge, and follow-up protocols.
- The Standing Committee reviewed the overall 25th to 75th percentile performance range of risk-standardized hospital visit rates per 1000 colonoscopies from 11.8% to 12.8%, with mean performance of 12.3% and agreed that there is a narrow performance gap.

---

2. Scientific Acceptability of Measure Properties: <u>The Standing Committee did not vote as the measure was withdrawn</u>

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

2a. Reliability: **H-X**; **M-X**; **L-X**; **I-X**; 2b. Validity: **H-X**; **M-X**; **L-X**; **I-X**

Rationale:

- The Standing Committee reviewed the reliability testing and noted differences in the measure specifications (one year of data) and the three years of simulated signal-to-noise testing data. The Standing Committee requested that the developer consider aligning the testing with the measure specifications and resubmit the measure using a three year time frame for consideration in a future cycle of measure review.
- The Standing Committee noted that this alignment would help facilitate transparency and understanding among stakeholders and those being measured.
- The developer agreed to withdraw the measure from consideration and resubmit in a future cycle with testing data and measure specifications using a three year time frame. Given the withdrawal from the process, voting on the measure was suspended.
- The measure was withdrawn from consideration and will be updated and submitted in a future measure review cycle.

---

3. Feasibility: H-X; M-X; L-X; I-X: The Standing Committee did not vote as the measure was withdrawn

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

---

4. Usability and Use: The Standing Committee did not vote as the measure was withdrawn

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-X; No Pass-**X; 4b. Usability: **H-X; M-X; L-X; I-X**

---

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Yes-X**; **No-X**

Rationale:

---

6. Public and Member Comment

NQF did not receive comments following the developer's withdrawal of the measure.

---

7. Consensus Standards Approval Committee (CSAC) Endorsement Decision: N/A

The CSAC did not vote on measure 2539 as it was withdrawn by the developer during the Committee evaluation period, prior to an endorsement recommendation. During the October 21 meeting and review of this project, CSAC members did have questions about 2539, noting that it is currently one of the only measures in use in both ASCs and outpatient quality reporting programs, and that it is being used at the one-year reporting level, which is valuable to consumers. The CSAC member asked why the Committee did not ask the developer to update the specifications to match the testing, rather than not recommending it. Committee co-chair Dr. Bulger noted that the measure was withdrawn by the

developer before they could make an endorsement recommendation. The developer was on the line and explained to the CSAC that during the time they were submitting the measure to NQF, CMS changed the specifications to a three-year measure. The reliability results are much stronger at a three-year level and that is the version that will be reported publicly in the near future, so the developer elected to withdraw the measure until they can update it. CSAC agreed with this rationale but noted the importance of more frequent reporting for consumers. CSAC members encouraged CMS to consider this in the future, noting the extremely limited information about safety and quality in outpatient settings; further, they noted that annual results are much more helpful for facilities themselves to help improve quality more rapidly.

---

8. Appeals

N/A

Appendix B: All Cause Admissions and Readmissions Portfolio—Use in Federal Programs[d]

NQF #	Title	Federal Programs: Finalized or Implemented as of November 7, 2019
0171	Acute Care Hospitalization During the First 60 Days of Home Health	Home Health Quality Reporting, Home Health Value Based Purchasing
0173	Emergency Department Use without Hospitalization During the First 60 Days of Home Health	Home Health Quality Reporting, Home Health Value Based Purchasing
0275	Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate (PQI 5)	Medicare Shared Savings Program, Medicaid
0277	Heart Failure Admission Rate (PQI 8)	Medicare Shared Savings Program, Medicaid
0330	Hospital 30-day, All Cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization	Hospital Inpatient Quality Reporting, Hospital Readmission Reduction Program
0505	Hospital 30-day All Cause risk-standardized readmission rate (RSRR) following acute myocardial infarction	Hospital Inpatient Quality Reporting, Hospital Readmission Reduction Program
0506	Hospital 30-day, All Cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization	Hospital Inpatient Quality Reporting, Hospital Readmission Reduction Program
0695	Hospital 30-Day Risk-Standardized	None

[d] Per CMS Measures Inventory Tool as of 12/9/2019

NQF #	Title	Federal Programs: Finalized or Implemented as of November 7, 2019
	Readmission Rates following Percutaneous Coronary Intervention (PCI)	
1551	Hospital-level 30-day, All Cause risk-standardized readmission rate (RSRR) following elective primary total hip arthroplasty (THA) and/or total knee arthroplasty (TKA)	Hospital Inpatient Quality Reporting, Hospital Readmission Reduction Program, Hospital Compare
1891	Hospital 30-Day, All Cause, Risk-Standardized Readmission Rate (RSRR) following Chronic Obstructive Pulmonary Disease (COPD)	Hospital Inpatient Quality Reporting, Hospital Readmission Reduction Program, Hospital Compare
1768	Plan All Cause Readmissions (PCR)	Medicare Part C Star Rating, Medicaid, Qualified Health Plan (QHP) Quality Rating System (QRS)
1789	Hospital-Wide All Cause Unplanned Readmission Measure (HWR)	Hospital Inpatient Quality Reporting, Medicare Shared Savings Program, Hospital Compare, Merit-Based Incentive Payment System (MIPS) Program
2375	PointRight ® Pro 30™	None
2393	Pediatric All-Condition Readmission Measure	None
2380	Rehospitalization During the First 30 Days of Home Health	Home Health Quality Reporting
2414	Pediatric Lower Respiratory Infection Readmission Measure	None
2496	Standardized Readmission Ratio	Dialysis Facility Compare, End Stage Renal Disease-Quality Incentive Program
2502	All Cause Unplanned Readmission Measure for 30 Days Post Discharge from	Inpatient Rehabilitation Facility Quality Reporting

NQF #	Title	Federal Programs: Finalized or Implemented as of November 7, 2019
	Inpatient Rehabilitation Facilities (IRF)	
2503	Hospitalizations per 1000 Medicare fee-for-service (FFS) Beneficiaries	None
2504	30-day Rehospitalizations per 1000 Medicare fee-for-service (FFS) Beneficiaries	None
2505	Emergency Department Use without Hospital Readmission During the First 30 Days of Home Health	Home Health Quality Reporting
2510	Skilled Nursing Facility 30-Day All Cause Readmission Measure	Skilled Nursing Facility Value-Based Purchasing
2512	30-Day All Cause Post Long-Term Care Hospital (LTCH) Discharge Hospital Readmission Measure	Long-term Care Hospital Quality Reporting
2513	Hospital 30-Day All Cause Risk-Standardized Readmission Rate (RSRR) following Vascular Procedures	None
2514	Risk-Adjusted Coronary Artery Bypass Graft (CABG) Readmission Rate	None
2515	Hospital 30-day, All Cause, unplanned, risk-standardized readmission rate (RSRR) following coronary artery bypass graft (CABG) surgery	Hospital Inpatient Quality Reporting, Hospital Readmission Reduction Program, Hospital Compare
2539	Facility 7-Day Risk-Standardized Hospital	Ambulatory Surgical Center Quality Reporting, Hospital Compare, Hospital Outpatient Quality Reporting

NQF #	Title	Federal Programs: Finalized or Implemented as of November 7, 2019
	Visit Rate after Outpatient Colonoscopy	
2827	PointRight® Pro Long Stay(TM) Hospitalization Measure	None
2858	Discharge to Community	None
2860	Thirty-day All Cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF)	Inpatient Psychiatric Facility Quality Reporting
2879	Hybrid Hospital-Wide Readmission (HWR) Measure with Claims and Electronic Health Record Data	Hospital Compare, Hospital Inpatient Quality Reporting
2880	Excess days in acute care (EDAC) after hospitalization for heart failure (HF)	Hospital Compare, Hospital Inpatient Quality Reporting
2881	Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI)	Hospital Compare, Hospital Inpatient Quality Reporting
2882	Excess days in acute care (EDAC) after hospitalization for pneumonia	Hospital Compare, Hospital Inpatient Quality Reporting
2886	Risk-Standardized Acute Admission Rates for Patients with Heart Failure	Medicare Shared Savings Program
2887	Risk-Standardized Acute Admission Rates for Patients with Diabetes	Medicare Shared Savings Program
2888	Risk-Standardized Acute Admission Rates	Medicare Shared Savings Program

NQF #	Title	Federal Programs: Finalized or Implemented as of November 7, 2019
	for Patients with Multiple Chronic Conditions	
3188	30-Day Unplanned Readmissions for Cancer Patients	None

Appendix C: All Cause Admissions and Readmissions Standing Committee and NQF Staff

STANDING COMMITTEE

John Bulger, DO, MBA (Co-chair)
Chief Quality Officer, Geisinger Health System
Danville, Pennsylvania

Cristie Travis, MSHHA (Co-chair)
Chief Executive Officer, Memphis Business Group on Health
Memphis, Tennessee

Katherine Auger, MD, MSc
Assistant Professor of Pediatrics, Cincinnati Children's Hospital Medical Center
Cincinnati, Ohio

Jo Ann Brooks, PhD, RN
Vice President of Safety and Quality, Indiana University Health System
Indianapolis, Indiana

Helen Chen, MD
Chief Medical Officer, Hebrew SeniorLife
Boston, Massachusetts

Susan Craft, RN
Director, Care Coordination Initiatives - Office of Clinical Quality & Safety, Henry Ford Health System
Detroit, Michigan

William Wesley Fields, MD, FACEP
Assistant Clinical Professor, UC Irvine Medical Center; Board of Directors, CEP America
Laguna Niguel, California

Steven Fishbane, MD
Chief Division of Kidney Diseases and Hypertension and Vice President, North Shore-LIJ Health System for Network Dialysis Services
Commack, New York

Laurent Glance, MD
Vice-Chair for Research, University of Rochester School of Medicine
Rochester, New York

Anthony Grigonis, PhD
Vice President, Quality Improvement, Select Medical
Mechanicsburg, Pennsylvania

Bruce Hall, MD, PhD, MBA
Professor, Surgeon, Washington University; Vice President for Patient Outcomes, BJC Healthcare
Saint Louis, Missouri

Leslie Kelly Hall
SVP Policy, Healthwise
Boise, Idaho

Paul Heidenreich, MD, MS, FACC, FAHA
Professor and Vice-Chair for Clinical, Quality, and Analytics, Stanford University School of Medicine, and VA Palo Alto Health Care System
Palo Alto, California

Karen Joynt Maddox, MD, MPH
Assistant Professor of Medicine, Washington University School of Medicine
St. Louis, MO

Keith Lind, JD, MS, BSN
Senior Policy Advisor, AARP Public Policy Institute
Washington, DC

Paulette Niewczyk, PhD, MPH
Director of Research, Uniform Data System for Medical Rehabilitation
Amherst, New York

Carol Raphael, MPA
Senior Advisor, Manatt Health Solutions
New York, New York

Mathew Reidhead, MA
Vice President of Research and Analytics, Missouri Hospital Association, Hospital Industry Data Institute
Jefferson City, Missouri

Pamela Roberts, PhD, MSHA, ORT/L, SCFES, FAOTA, CPHQ
Manager for Inpatient Rehabilitation; Quality, Education, and Research; and Neuropsychology, Cedars-Sinai Medical Center
Los Angeles, California

Derek Robinson, MD, MBA, FACEP, CHCQM
Vice President for Quality and Accreditation, Health Care Service Corporation
Chicago, Illinois

Thomas Smith, MD, FAPA
Medical Director, Division of Managed Care, NYS Office of Mental Health, Special Lecturer, Columbia University Medical Center
New York, New York

NQF STAFF

Elisa Munthali, MPH
Senior Vice President, Quality Measurement

Andrew Lyzenga, MPP
Senior Director, Quality Measurement

Suzanne Theberge, MPH
Senior Project Manager

Oroma Igwe, MPH
Project Manager

Asaba Mbenwoh Nguafor, RN, MSN/MPH
Project Analyst

Taroon Amin, PhD, MPH
NQF Consultant

Appendix D: Pre-Evaluation Comments

Comments received as of June 12, 2019.

2539 Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy

Submitted by American Medical Association (AMA)

The American Medical Association (AMA) appreciates the opportunity to comment on this measure. We are writing to express our concerns on the evidence and testing provided in support of this measure.

While the AMA agrees that it is useful to understand the rate of complications following outpatient colonoscopies for quality improvement, we did not see explicit information outlining how these facilities can implement structures or processes that can lead to improved outcomes for these patients. Rather, most of the cited references focused on incidence rates and prevalence of specific risk factors and did not address what factors or processes leveraged by a facility can reduce the occurrence of complications.

Regarding the validity of the measure and specifically the risk adjustment approach, we do not believe that the measure is adequately tested and adjusted for social risk factors. The conceptual basis for the selection of the social risk factors was inadequately described in section 2.b.3. Risk Adjustment/Stratification and it is unclear to us why the developer would test social risk factors after adjusting for clinical risk factors rather than assessing the impact of both clinical and social risk factors in the model at the same time. These variations in how risk adjustment factors are examined could also impact how each variable (clinical or social) perform in the model and remain unanswered questions.

In addition, the AMA questions whether the information provided as a result of this measure is truly useful for accountability and informing patients of the quality of care provided by hospital outpatient departments (HOPDs) or ambulatory surgical centers (ASCs). Specifically, our concern relates to the relatively limited amount of variation across applicable facilities. Only two HOPDs out of the 3,908 facilities were identified as performing “Better than the National Rate” or “Worse than the National Rate” and of 2,061 ASCs, none were identified as performing “Better than the National Rate” and four performed “Worse than the National Rate.” Endorsing a measure that currently only identifies a small number of outliers does not enable users to distinguish meaningful differences in performance and is inconsistent with the validity subcriterion and usability/use criterion.

We ask the Standing Committee to carefully consider these concerns during their evaluation.

3495 Hospital-Wide 30-Day, All Cause, Unplanned Readmission (HWR) Rate for the Merit-Based Incentive Payment System (MIPS) Eligible Clinicians and Eligible Clinician Groups: Change in Mobility Score

Submitted by American Medical Association (AMA)

The American Medication Association (AMA) appreciates the opportunity to comment on this measure. Below we outline our concerns on whether this measure meets the NQF Measure Evaluation Criteria, particularly for evidence and scientific acceptability.

The AMA believes that attribution must be determined based on evidence that the accountable unit is able to meaningfully influence the outcome, which aligns with the most recent National Quality Forum (NQF) report, Improving Attribution Models (NQF, 2018). This principle is also aligned with the evidence

requirements for outcome measures in the NQF Measure Evaluation Criteria, which requires that there be at least one structure or process that can influence the outcome and this relationship must be demonstrated through empirical evidence. CMS must begin to demonstrate these relationships with the accountable unit prior to implementing this measure in MIPS and we do not believe that CMS has adequately demonstrated this link.

While the AMA agrees that there is evidence to demonstrate that improved care coordination and programs focused on discharge planning can lead to reductions in hospital readmissions, most of the cited evidence involved multiple partners and clinicians such as the health system, hospital, nurse, and/or pharmacist. Therefore, we do not believe that sufficient evidence was provided to support that physicians or practices using the proposed attribution approach in the absence of some coordinated program or targeted intervention led by the health system or hospital can implement structures or processes leading to improved outcomes for these patients.

In addition, continuity of care requires smooth transitions to prepare for patients' changing clinical and social needs, but the Stark law often impedes the continuity and care transitions. Specifically, in certain circumstances, physicians are prohibited from employing promising care coordination strategies on behalf of their patients, e.g., an arrangement that pays for a nurse coordinator to coordinate a recently discharged patient's care among a hospital, physician specialists, or a primary care physician due to concerns that this may induce future referrals to their own office to avoid an unnecessary readmission to the hospital. As a result, we do not believe that assignment of responsibility of the reduction of readmissions to multiple physicians and practices in MIPS is appropriate nor has the developer provided sufficient information to support the attribution of this measure to up to three physicians or practices.

The AMA is disappointed to see the low measure score reliability results based on the minimum case number of 25 patients. We believe that measures must meet minimum acceptable thresholds of 0.7 for reliability and during the public comment period in December 2018, the data provided when using at least 100 patients yielded mean signal-to-noise results of 0.991 for eligible clinicians and 0.997 for eligible clinician groups (CMS, 2018). We request that the Standing Committee evaluate whether the case minimum of 25 patients is acceptable given the low reliability results.

The AMA is also troubled to see that no evidence or testing has been provided to support the attribution of this measure to the three distinct groups (discharge physician, primary inpatient care provider, and outpatient primary care provider). While correlations to the hospital's overall star ratings and readmission score from the star ratings are useful, we do not believe that the developer has provided sufficient information as it relates to the measure's application to each of the accountable units to which the measure is attributed.

In addition, we noted that the conceptual basis used to explain which social risk factors were tested in Section 2b3.3a solely focused on the hospital and was not specific to physicians or practices. It is difficult to determine whether additional factors should be considered without this information and we do not believe that it is responsive to NQF criteria requirements.

We also remain concerned that CMS continues to test social risk factors after assessment of clinical and demographic risk factors and it is unclear why this multi-step approach is preferable. On review of the Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors report, it is clear that the approaches to testing these data should be revised to strategies such as multi-level models or testing of social factors prior to clinical factors and that as access to new data becomes available, it may elucidate more differences that are unrelated to factors within a hospital's or physician's control (NQF, 2017). Additional testing is needed to evaluate clinical factors in-conjunction with social risk factors; as opposed to the current approach that prioritizes clinical factors. Even though the c-statistics for each cohort were not improved, it would be useful to understand further the impact that the inclusion of

these factors had on the absolute change of the rates since the differences ranged from a minimum of -1.13% to a maximum of 3.99% for eligible clinicians and a minimum of -2.88% to a maximum of 4.24% for eligible clinician groups. These shifts could potentially impact the points physicians score in the Quality Category in MIPS and as a result, either positively or negatively impact the overall penalty or incentive they receive and the resources available for those individuals and groups who serve larger numbers of disadvantaged patients.

Given the measure is specifically developed for MIPS, the developer must perform testing that demonstrates how the measure would perform under the MIPS benchmark methodology and Physician Compare Star Ratings since CMS utilizes two different methodologies for ranking and profiling physicians.

In conclusion, CMS must balance the desire to apply these measures to the broadest number of clinicians possible with the unintended consequences of inappropriately attributing measures to physicians for which they cannot meaningfully influence patient outcomes. The AMA requests that the Standing Committee carefully consider the potential misinformation that could be provided to patients and caregivers if the measures do not have a clear evidence base to support attribution of the outcome to a specific physician and could potentially produce scores that are invalid and unreliable.

References:

Centers for Medicare and Medicaid Services. Development and Reevaluation of Outpatient Outcome Measures for the Merit-based Incentive Payment System. Public Comment Page: Currently Accepting Comments. Available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Outpatient-MIPS-MCC-Measure-Development.zip. Last accessed June 11, 2019.

http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=88154National Quality Forum. Improving Attribution Models. Final Report. August 31, 2018. Available at: http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=88154. Last accessed December 18, 2018.

National Quality Forum. Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors. Final report. July 18, 2017. Available at: http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=85635. Last accessed December 18, 2018.

National Quality Forum
1030 15th St NW, Suite 800
Washington, DC 20005
http://www.qualityforum.org

ISBN