TO: Consensus Standards Approval Committee (CSAC)

FR: NQF Staff

RE: All-Cause Admissions and Readmissions 2015-2017 Project

DA: November 15, 2016

The CSAC will review recommendations from the Readmissions 2015-2017 project at its November 9-10, 2016 meeting.

This memo includes a summary of the project, recommended measures, and themes identified from and responses to the public and member comments.

Accompanying this memo are the following documents:

1. **Readmissions Draft Report.** The draft report has been updated to reflect the changes made following Standing Committee discussion of public and member comments. The complete draft report and supplemental materials are available on the project page.
2. **Comment Table.** Staff has identified themes within the comments received. This table lists 60 comments received during the post meeting comment period and the NQF/Standing Committee responses.

The Admissions/Readmissions Standing Committee has also been meeting to review 14 measures that were endorsed with conditions. Details of that review are included in a separate memo.

BACKGROUND

Reducing avoidable admissions and readmissions to acute care facilities continues to be an important focus of quality improvement across the healthcare system. Unnecessary hospitalizations can prolong a patient's illness, increase their time away from home and family, expose them to potential harms, and add to their costs. Avoidable admissions and readmissions also contribute significantly to the United States' high rate of healthcare spending. One estimate puts the cost of all-cause adult hospital readmissions at over 40 billion dollars annually. While there is no clear evidence on how many of these readmissions may be avoidable, estimates have ranged that anywhere from five percent to 79 percent may be preventable.[i] A 2013 MedPAC report suggests that reducing avoidable readmissions by 10 percent could achieve a savings of $1 billion or more.[ii]

Currently, there are more than 46 NQF-endorsed admissions and readmissions. These measures have been adopted into a number of federal quality programs with the aim of reducing unnecessary admissions and readmissions by fostering improved care coordination across the healthcare system.

DRAFT REPORT

The All-Cause Admissions and Readmissions 2015-2017 Draft Report presents the results of the evaluation of 17 measures considered under the Consensus Development Process (CDP). The Standing Committee evaluated 11 newly-submitted measures and 6 measures undergoing maintenance review against NQF's

standard evaluation criteria. Sixteen measures were recommended for endorsement.

The measures were evaluated against the 2016 version of the measure evaluation criteria.

	Maintenance	New	Total
Measures under consideration	6	11	17
Measures recommended for endorsement	6	10	16
Measures not recommended for endorsement	0	1	1
Reasons for not recommending	Importance – 0 Scientific Acceptability – 0 Overall – 0 Competing Measure – 0	Importance – 0 Scientific Acceptability – 1 Overall – 0 Competing Measure – 0	

All-Cause Admission and Readmission Measures Recommended for Endorsement:

- #0171 Acute Care Hospitalization During the First 60 Days of Home Health (CMS)
  Overall Suitability for Endorsement: Y-19; N-0
- #0173 Emergency Department Use without Hospitalization During the First 60 Days of Home Health (CMS)
  Overall Suitability for Endorsement: Y-16; N-0
- #0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization (Yale/CORE)
  Overall Suitability for Endorsement: Y-19; N-1
- #0506 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization (Yale/CORE)
  Overall Suitability for Endorsement: Y-18; N-1
- #1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR) (Yale/CORE)
  Overall Suitability for Endorsement: Y-18; N-1
- #1891 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization (Yale/CORE)
  Overall Suitability for Endorsement: Y-18; N-2
- #2827 PointRight® Pro Long Stay(TM) Hospitalization Measure (PointRight)
  Overall Suitability for Endorsement: Y-18; N-1
- #2858 Discharge to Community (American Health Care Association)
  Overall Suitability for Endorsement: Y-19; N-0
- #2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF) (Health Services Advisory Group, Inc.)
  Overall Suitability for Endorsement: Y-19; N-1
- #2879 Hybrid Hospital-Wide Readmission Measure with Claims and Electronic Health Record Data (Yale/CORE)
  Overall Suitability for Endorsement: Y-16; N-2
- #2880 Excess days in acute care (EDAC) after hospitalization for heart failure (Yale/CORE)

Overall Suitability for Endorsement: Y-16; N-1

- #2881 Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI) (Yale/CORE)
  Overall Suitability for Endorsement: Y-18; N-0
- #2882 Excess days in acute care (EDAC) after hospitalization for pneumonia (Yale/CORE)
  Overall Suitability for Endorsement: Y-18; N-0
- #2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure (Yale/CORE)
  Overall Suitability for Endorsement: Y-19; N-1
- #2887 Risk-Standardized Acute Admission Rates for Patients with Diabetes (Yale/CORE)
  Overall Suitability for Endorsement: Y-18; N-2
- #2888 Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions (Yale/CORE)
  Overall Suitability for Endorsement: Y-20; N-0

Readmission Measure Not Recommended (See Appendix A for the Committee's votes and rationale)

- #2884 30-Day Unplanned Readmissions for Cancer Patients (ADCC)

COMMENTS AND THEIR DISPOSITION

NQF received 60 comments from twenty-eight member organizations pertaining to the general draft report and to the measures under consideration.

A table of comments submitted during the comment period, with the responses to each comment and the actions taken by the Standing Committee and measure developers, is posted to the All-Cause Admissions and Readmission 2015-2017 project page under the Public and Member Comment section.

Commenters were asked for input on both the new project and the measures endorsed with conditions. Themes around the consideration of sociodemographic factors and the Committee's response apply to measures in both the 2015-2017 project and the review of measures endorsed with conditions.

Comment Themes and Committee Responses

Comments about specific measure specifications and rationale were forwarded to the developers, who were invited to respond.

The Standing Committee reviewed all of the submitted comments and developer responses. Committee members focused their discussion on measures or topic areas with the most significant and recurring issues.

Theme 1 – Consideration of Sociodemographic Status Factors

Many commenters expressed concern regarding potentially insufficient adjustments made for sociodemographic status (SDS) factors. The comments submitted to NQF urged the Committee to take a more in-depth look at the need for SDS adjustment, given the potentially negative impact these measures could have on providers practicing in low-resource regions. Some commenters noted that the findings presented by measure developers who did not include these factors in their measure contradict common knowledge and findings from other research. Commenters encouraged additional testing of SDS factors and stratifying measure results by SDS factors such as dual eligibility for Medicare and Medicaid.

Committee Response: The Committee has reviewed your comment and appreciates your input.

Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the Committee agreed with the developer's decision not to include these factors at this time, the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area. More work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

Developer Response: Concerning the issue of using race as a proxy for socioeconomic status (SES), we agree with the AHA and with the NQF's guidance suggesting that race should not be used as a proxy for SES. Race was not used in the analyses as a proxy for SES but as an important comparator with SES variables. Although the NQF Expert Panel on Risk Adjustment for Sociodemographic (SDS) Factors did not provide clear guidance regarding the inclusion of race in measure's risk models, the panel did broaden the term from SES to SDS to account for consideration of racial disparities, and we feel it is useful to understand the pattern of racial disparities along with SES disparities. Therefore, we believe it is helpful to show analyses with race, not because it should be included in risk-adjustment models, but as a point of comparison with SES variables. The conceptual rationale for not including SES variables in the measures' risk models has important parallels with race in that both SES and race are associated with access to differential quality hospitals and can lead to differential care within hospitals. These comparisons can be helpful in understanding causal pathways and for making decisions about incorporating SES variables in risk-adjustment models.

Theme 2 – Level-of-Analysis & Implementation

Commenters raised concerns about the use of NQF-endorsed measures at a different level of analysis than the one for which they are endorsed. In particular, a number of commenters raised concerns that NQF #1789 *Hospital-wide all-cause unplanned readmission measure* is being used at the clinician level of analysis in the Physician Value-Based Payment Modifier program and is proposed to be used in the Merit-Based Incentive Payment System in a similar way. These commenters expressed concern that testing at this level of analysis was not provided to the Standing Committee for review. Commenters expressed concerns that other measures could also be used at a different level of analysis than the one for which they are endorsed.

NQF Response: Thank you for your comment. NQF endorses measures specifically for the level of analysis indicated in the measure specifications. Additionally, the level of analysis must be support by reliability and validity testing.

Committee Response: Thank you for your comment. The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.

Theme 3 – Data Limitations

Commenters raised some particular concerns about applying measures that incorporate electronic clinical data at the health plan level.

Committee Response: The Committee has reviewed your comment and appreciates your input. The Committee agrees that the measure should be applied at the facility-level, as it is specified and tested. The Committee believes that linking claims and EHR data is an important advancement in quality measurement.

Theme 4 - Potentially Competing Measures

One commenter expressed concern that the current NQF portfolio of readmission measures contains unnecessary overlap in condition or setting assessment. The commenter urged the Committee to select "best in class" measures and implored NQF to facilitate opportunities to do so.

Committee Response: The Committee followed NQF's guidance on measure harmonization throughout the evaluation process. Prior to the in-person meeting, the Committee received materials regarding these competing measures, and held a separate call after the in-person meeting on September 1 to discuss harmonization issues and allow the developers to answer questions from Committee members. The Committee then voted via survey to recommend both measures. The Committee considered the added value and burden of recommending both measures and agreed that the differences in measure specifications added sufficient value to offset any potential negative impact.

Theme 5 - Potential Negative Unintended Consequences

Commenters raised a number of concerns related to potential negative unintended consequences of the use of readmissions measures. Commenters noted the inverse correlation between readmissions and mortality. Commenters also raised concerns about the relationship between decreasing admission rates and the readmission measures.

Committee Response: The Committee has reviewed your comment and appreciates your input. The Committee recognizes the potential for negative unintended consequences of admissions and readmissions

measures and recommends careful monitoring of their implementation. Above all, the Committee agreed that use of these measures should be monitored to ensure they do not inadvertently reduce access to necessary care. The Committee noted the inverse relationship between mortality and readmission for heart failure and recognized the need for careful surveillance and balancing of these measures. The Committee also reiterated its concerns about the need to carefully balance implementation of measures addressing psychiatric readmissions to prevent the risk of higher suicidality.

On the other hand, the Committee has noted the desire to understand a patient's need for any subsequent acute care after a hospitalization. In particular, the Committee recognized the need understand if patients are being seen in the Emergency Department after discharge or being placed in observation. The Committee recommends continued work to ensure that the use of readmissions measures does not result in unnecessary or avoidable use of the emergency department or observation status while ensuring that all patients have access to any necessary care. The Committee noted that a number of measures recommended for endorsement in this project could help to balance these concerns, in particular the measures addressing excess days in acute care and population-based admission measures.

NQF MEMBER VOTING RESULTS

One of the recommended measures were approved with 67% approval or higher. Representatives of 19 member organizations voted; no votes were received from Consumer, Supplier/Industry, or Public/Community Health Agency Councils. Results for each measure are provided in Appendix B.

Appendix A – Measures Not Recommended for Endorsement

The table below lists the Committee's vote and rationale for measures not recommended for endorsement.

Measure	Voting Results	**Rationale:**
#2884 30-Day Unplanned Readmissions for Cancer Patients	**Initial Vote:** **Evidence** Y-17; N-0 **Gap** H-1; M-16; L-0; I-1 **Reliability** H-0; M-5; L-13; I-1	The Standing Committee recommended that the developers separate out payer class as a marker of socioeconomic challenges. In particular, the Standing Committee raised concerns about the unique challenges Medicaid patients face when seeking treatment for cancer and recommended that they not be categorized with patients who are opting to pay for treatment out of pocket. The Standing Committee also suggested the developer consider ways to track readmissions to other facilities and to consider a longer time window.

Appendix B – NQF Member Voting Results

NQF MEMBER VOTING RESULTS

One of the recommended measures were approved with 67% approval or higher. Representatives of 19 member organizations voted; no votes were received from Consumer, Supplier/Industry, or Public/Community Health Agency Councils. Results for each measure are provided below.

NQF Member Council	Voting	Eligible to Vote	Rate
Consumer	0	40	0%
Health Plan	1	18	6%
Health Professional	8	103	8%
Provider Organizations	5	108	5%
Public/Community Health Agency	0	17	0%
Purchaser	2	21	10%
QMRI	3	77	4%
Supplier/Industry	0	36	0%
All Councils	**19**	**420**	**4%**

Measure #0171 Acute Care Hospitalization During the First 60 Days of Home Health (CMS)

Measure Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	1	1	6	8	50%
Provider Organizations	1	3	1	5	25%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%
Supplier/Industry	0	0	0	0	
All Councils	7	4	8	19	**64%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					75%

*equation: Yes/ (Total - Abstain)

Voting Comments

Federation of American Hospitals: The FAH has serious concerns about the potential endorsement of this set of measures as this point in time. While the FAH appreciates that NQF initiated a trial process for socio-demographic status (SDS) adjustment and the results of that work are not yet complete, the FAH is concerned that the SDS review process has raised more questions than it has answered. The measures being considered in this readmission project are a good example.
The FAH will submit full comments in a separate letter. In general, the FAH believes the Project would benefit from overall review of the measures to determine which factors are best for risk adjustment. Also, the Committee should review the measures to evaluate the appropriateness of methodology, i.e., all-cause readmission versus preventable readmission. The Committee should also question/explore the need for a readmission measure for a large number of clinical conditions. Are they providing patients with useful information to make informed choices?

For instance, the FAH still does not believe that a strong conceptual basis has been established to determine/facilitate the gathering and evaluation of empirical evidence for SDS adjustment. This lack of a strong framework for consideration must be addressed before NQF proceeds with a final decision on the measures in this project.

The FAH does not believe that the conceptual model, associated risk variables and results were adequately discussed for these measures. Based on the summary of the Committee's evaluation in the report, it appears that they had a very limited discussion on the new information on the consideration of SDS factors in the risk model. The Committee appears to have only focused on the results of the developer's analysis. The FAH does not believe that the Committee had a meaningful discussion of the conceptual model for risk adjustment presented by the developer. In addition, there was little evaluation of whether the SDS risk factors used in the measures under review were correlated sufficiently to the model.
The FAH will submit additional comments in a separate letter sent to the Committee. The FAH comments apply to all of the measures in this project.

Adventist Health System: Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk. We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves

a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum.

Measure #0173 Emergency Department Use without Hospitalization During the First 60 Days of Home Health (CMS)

Measure Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	1	1	6	8	50%
Provider Organizations	2	2	1	5	50%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%
Supplier/Industry	0	0	0	0	
All Councils	8	3	8	19	**73%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					80%

*equation: Yes/ (Total - Abstain)

Measure #0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization (Yale/CORE)

Measure Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	2	6	8	0%
Provider Organizations	1	4	0	5	20%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	1	0	2	3	100%
Supplier/Industry	0	0	0	0	
All Councils	5	6	8	19	**45%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					64%

*equation: Yes/ (Total - Abstain)

Voting Comments

Adventist Health System: Adventist Health System appreciates the Committee and the developer's responses to the concerns raised by commenters. However, we do not feel as though the issues raised were adequately addressed. Our organization commends the Committee for stressing "the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk.

We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum.

AHS is thankful for the developer's thoughtful response to our concerns about this measure's Inter-Class Correlation Coefficient (ICC) and C-statistic. We understand developer's rationale for its approach to assessing reliability and commend it for using such a rigorous method. In addition, we appreciate the developer's citation of a convention that it believes is "more appropriate." As the developer has stated, that convention "describes the ICC values as moderate (0.41-0.60) for this measure." We agree with Landis and Koch [1977] that "[a]though these divisions are clearly arbitrary, they do provide useful 'benchmarks' for the discussion of [a] specific example […]" (Landis JR and Koch GG. The Measurement of Observer Agreement for Categorical Data.
Biometrics 1977; 33:159-174). Furthermore, we agree with the developer that the ICC values of this measure could be described as "moderate" under the "benchmarks" put forward by Landis and Koch [1977].

However, AHS believes that NQF endorsement should only be awarded to measures if they meet a threshold of reliability commensurate with the impact of their use. It is our opinion that achieving a "moderate" benchmark of reliability is it believes is "more appropriate." As the developer has stated, that convention "describes the ICC values as moderate (0.41-0.60) for this measure." We agree with Landis and

Koch [1977] that "[a]though these divisions are clearly arbitrary, they do provide useful 'benchmarks' for the discussion of [a] specific example […]" (Landis JR and Koch GG. The Measurement of Observer Agreement for Categorical Data.

Biometrics 1977; 33:159-174). Furthermore, we agree with the developer that the ICC values of this measure could be described as "moderate" under the "benchmarks" put forward by Landis and Koch [1977].

However, AHS believes that NQF endorsement should only be awarded to measures if they meet a threshold of reliability commensurate with the impact of their use. It is our opinion that achieving a "moderate" benchmark of reliability is not a sufficient for the endorsement of substantially impactful measures. We find measures that are used in public reporting or payment programs, such as the Hospital Inpatient Quality Reporting Program (IQR) and Hospital Readmissions Reduction Program (HRRP), to be substantially impactful. Hence, AHS believes that for these measures to be awarded endorsement they should first be assessed as meeting a reliability benchmark or "strength of agreement" that is "substantial" according to Landis and Koch [1977]. Thus, we have concluded that, according to the "more appropriate convention" described by the developer, the "substantial" reliability "benchmark" for this measure, as outlined by Landis and Koch [1977], would be an ICC value of 0.61-0.80. Therefore, the assessed level of reliability of this measure is another reason why we are voting no.

Measure #0506 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization (Yale/CORE)

Measure Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	2	6	8	0%
Provider Organizations	1	4	0	5	20%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%
Supplier/Industry	0	0	0	0	
All Councils	6	6	7	19	**50%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					64%

*equation: Yes/ (Total - Abstain)

Voting Comments

Adventist Health System: Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we

are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk. We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum. In addition, we remain concerned about the broadening of the measure cohort.

Measure #1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR) (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	0	8	8	
Provider Organizations	1	4	0	5	20%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	1	1	1	3	50%
Supplier/Industry	0	0	0	0	
All Councils	5	5	9	19	**50%**
Percentage of councils approving (>60%)					50%
Average council percentage approval					68%

*equation: Yes/ (Total - Abstain)

Voting Comments

University of Iowa Public Policy Center – Health Policy Research Program: We have attempted to use this measure with the Medicaid population and felt it was extremely restrictive and difficult to apply the unplanned portion.

American Urological Association: The AUA has concerns about the level of use--at the physician and group practice levels--and risk adjustment. We request that the measure developer provide evidence

supporting the reliability and validity at the physician and group practice level.

Society of Hospital Medicine: SHM continues to have concerns about the use of this measure beyond its original intent through its incorporation into physician pay for performance programs.

Adventist Health System: Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk. We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum. We also remain concerned, as the Committee noted, that "merging multiple cohorts into one group may mask the individual variance properties of the individual cohorts." We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum. We also remain concerned, as the Committee noted, that "merging multiple cohorts into one group may mask the individual variance properties of the individual cohorts."

Measure #1891 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	2	6	8	0%
Provider Organizations	1	4	0	5	20%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	1	0	2	3	100%
Supplier/Industry	0	0	0	0	
All Councils	5	6	8	19	**45%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					64%

*equation: Yes/ (Total - Abstain)

Voting Comments

Adventist Health System appreciates the Committee and the developer's responses to the concerns raised by commenters. However, we do not feel as though the issues raised were adequately addressed. and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk.

We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference

between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum.

AHS is thankful for the developer's thoughtful response to our concerns about this measure's Inter-Class Correlation Coefficient (ICC) and C-statistic. We understand developer's rationale for its approach to assessing reliability and commend it for using such a rigorous method. In addition, we appreciate the developer's citation of a convention that it believes is "more appropriate." As the developer has stated, that convention "describes the ICC values as moderate (0.41-0.60) for this measure." We agree with Landis and Koch [1977] that "[a]though these divisions are clearly arbitrary, they do provide useful 'benchmarks' for the discussion of [a] specific example […]" (Landis JR and Koch GG. The Measurement of Observer Agreement for Categorical Data.
Biometrics 1977; 33:159-174). Furthermore, we agree with the developer that the ICC values of this measure could be described as "moderate" under the "benchmarks" put forward by Landis and Koch [1977].

However, AHS believes that NQF endorsement should only be awarded to measures if they meet a threshold of reliability commensurate with the impact of their use. It is our opinion that achieving a "moderate" benchmark of reliability is not a sufficient for the endorsement of substantially impactful measures. We find measures that are used in public reporting or payment programs, such as the Hospital Inpatient Quality Reporting Program (IQR) and Hospital Readmissions Reduction Program (HRRP), to be substantially impactful. Hence, AHS believes that for these measures to be awarded endorsement they should first be assessed as meeting a reliability benchmark or "strength of agreement" that is "substantial" according to Landis and Koch [1977]. Thus, we have concluded that, according to the "more appropriate convention" described by the developer, the "substantial" reliability "benchmark" for this measure, as outlined by Landis and Koch [1977], would be an ICC value of 0.61-0.80. Therefore, the assessed level of reliability of this measure is another reason why we are voting no.

Measure #2827 PointRight® Pro Long Stay(TM) Hospitalization Measure (PointRight)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	2	6	8	0%
Provider Organizations	1	2	2	5	33%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%
Supplier/Industry	0	0	0	0	
All Councils	6	4	9	19	**60%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					67%

*equation: Yes/ (Total - Abstain)

Measure #2858 Discharge to Community (American Health Care Association)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	1	1	6	8	50%
Provider Organizations	1	2	2	5	33%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	1	0	3	67%
Supplier/Industry	0	0	0	0	
All Councils	7	4	8	19	**64%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					70%

*equation: Yes/ (Total - Abstain)

Measure #2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF) (Health Services Advisory Group, Inc.)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	2	6	8	0%
Provider Organizations	1	2	2	5	33%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	1	1	1	3	50%
Supplier/Industry	0	0	0	0	
All Councils	5	5	9	19	**50%**
Percentage of councils approving (>60%)					40%
Average council percentage approval					57%

*equation: Yes/ (Total - Abstain)

Voting Comments

University of Iowa Public Policy Center - Health Policy Research Program: This measure is far too difficult to compare over areas and time.

Measure #2879 Hybrid Hospital-Wide Readmission Measure with Claims and Electronic Health Record Data (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	4	4	8	0%
Provider Organizations	1	3	1	5	25%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	3	0	0	3	100%
Supplier/Industry	0	0	0	0	
All Councils	7	7	5	19	**50%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					65%

*equation: Yes/ (Total - Abstain)

Voting Comments

American Urological Association: The AUA has concerns about this measure as it is a re-engineered version of measure 1789.

Federation of American Hospitals: The FAH is concerned that the proliferation of readmission measures using new constructs such as the hybrid measure here, does not adequately assess which measure construct really will assist hospitals to improve the delivery of patient care.

Measure #2880 Excess days in acute care (EDAC) after hospitalization for heart failure (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	3	5	8	0%
Provider Organizations	1	3	1	5	25%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%
Supplier/Industry	0	0	0	0	
All Councils	6	6	7	19	**50%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					65%

*equation: Yes/ (Total - Abstain)

Voting Comments

Society of Hospital Medicine: SHM does not support the Excess Days in Acute Care measures as they overlap significantly with existing readmission measures. We view these measures as penalizing providers for the same patients and cases.

Measure #2881 Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI) (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	3	5	8	0%
Provider Organizations	1	4	0	5	20%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%
Supplier/Industry	0	0	0	0	
All Councils	6	7	6	19	**46%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					64%

*equation: Yes/ (Total - Abstain)

Voting Comments

Society of Hospital Medicine: SHM does not support the Excess Days in Acute Care measures as they overlap significantly with existing readmission measures. We view these measures as penalizing providers for the same patients and cases.

Adventist Health System: Adventist Health System appreciates the Committee and the developer's responses to the concerns raised by commenters. However, we do not feel as though the issues raised were adequately addressed. Our organization commends the Committee for stressing "the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk.

We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in

providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum. AHS is thankful for the developer's thoughtful response to our concerns about this measure's Inter-Class Correlation Coefficient (ICC) and C-statistic. We understand developer's rationale for its approach to assessing reliability and commend it for using such a rigorous method. In addition, we appreciate the developer's citation of a convention that it believes is "more appropriate." As the developer has stated, that convention "describes the ICC values as moderate (0.41-0.60) for this measure." We agree with Landis and Koch [1977] that "[a]though these divisions are clearly arbitrary, they do provide useful 'benchmarks' for the discussion of [a] specific example […]" (Landis JR and Koch GG. The Measurement of Observer Agreement for Categorical Data.

Biometrics 1977; 33:159-174). Furthermore, we agree with the developer that the ICC values of this measure could be described as "moderate" under the "benchmarks" put forward by Landis and Koch [1977]. However, AHS believes that NQF endorsement should only be awarded to measures if they meet a threshold of reliability commensurate with the impact of their use. It is our opinion that achieving a "moderate" benchmark of reliability is not a sufficient for the endorsement of substantially impactful measures. We find measures that are used in public reporting or payment programs, such as the Hospital Inpatient Quality Reporting Program (IQR) and Hospital Readmissions Reduction Program (HRRP), to be substantially impactful. Hence, AHS believes that for these measures to be awarded endorsement they should first be assessed as meeting a reliability benchmark or "strength of agreement" that is "substantial" according to Landis and Koch [1977]. Thus, we have concluded that, according to the "more appropriate convention" described by the developer, the "substantial" reliability "benchmark" for this measure, as outlined by Landis and Koch [1977], would be an ICC value of 0.61-0.80. Therefore, the assessed level of reliability of this measure is another reason why we are voting no.

Measure #2882 Excess days in acute care (EDAC) after hospitalization for pneumonia (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	3	5	8	0%
Provider Organizations	1	3	1	5	25%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	3	0	0	3	100%
Supplier/Industry	0	0	0	0	
All Councils	7	6	6	19	**54%**

Percentage of councils approving (>60%)	60%
Average council percentage approval	65%

*equation: Yes/ (Total - Abstain)

Voting Comments

Society of Hospital Medicine: SHM does not support the Excess Days in Acute Care measures as they overlap significantly with existing readmission measures. We view these measures as penalizing providers for the same patients and cases.

Adventist Health System: Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk. We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum.

Measure # 2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	3	5	8	0%
Provider Organizations	1	3	1	5	25%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%

Supplier/Industry	0	0	0	0	
All Councils	6	6	7	19	**50%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					65%

*equation: Yes/ (Total - Abstain)

Voting Comments

Adventist Health System: Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk.

We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum. In addition, we remain concerned about the broadening of the measure cohort. Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk.

We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum.

Measure #2887 Risk-Standardized Acute Admission Rates for Patients with Diabetes (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	3	5	8	0%
Provider Organizations	1	3	1	5	25%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	3	0	0	3	100%
Supplier/Industry	0	0	0	0	
All Councils	7	6	6	19	**54%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					65%

*equation: Yes/ (Total - Abstain)

Voting Comments

Adventist Health System: Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk. We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will

elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum.

Measure #2888 Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions (Yale/CORE)

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	0	0	0	0	
Health Plan	1	0	0	1	100%
Health Professional	0	3	5	8	0%
Provider Organizations	1	3	1	5	25%
Public/Community Health Agency	0	0	0	0	
Purchaser	2	0	0	2	100%
QMRI	2	0	1	3	100%
Supplier/Industry	0	0	0	0	
All Councils	6	6	7	19	**50%**
Percentage of councils approving (>60%)					60%
Average council percentage approval					65%

*equation: Yes/ (Total - Abstain)

Voting Comments

Adventist Health System: Adventist Health System appreciates that the Committee has "stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards." Furthermore, we understand that the Committee's focus was on the adjustments the developer was able to put forward at this time and that the Committee was not comfortable with the level of significance that those adjustments reached. That said, we are voting no because we believe that rather than waiting for continued deliberations and future reevaluations, the time is now for the NQF to push the field forward by limiting endorsement until sufficient SDS variable data is made available and assessments can be concluded with high confidence. AHS finds it disconcerting that a measure might be endorsed despite an expert panel's emphasis of the existence of a "high risk of unintended consequences," such as the unfair penalization of facilities that serve vulnerable patient populations, due to inadequate adjustment for SDS risk. We also wish to restate our concern that there is a misconception regarding the impacts of SDS risk adjustment on quality measurement. AHS believes that proper adjustment for SDS variables will elucidate true differences in health care quality, will enable more fair comparisons of dissimilar facilities and will allow for better recognition of potential disparities that may be masked by differences in providers' patient populations. For instance, if hypothetically the vast majority of Provider A's patient population are relatively wealthy individuals, its quality measure performance may disproportionately reflect the care those socioeconomically fortunate patients receive and mask potential disparities of care that may exist among the small minority of socioeconomically vulnerable patients that Provider A serves. In contrast, if hypothetically Provider B serves a disproportionate share of socioeconomically vulnerable patients, it may perform worse on the quality measure than Provider A. However, when the measure is adjusted for, or even simply stratified by, the correct SDS variables it may be possible to illuminate true performance difference between Provider B and Provider A. In fact, Provider B may outperform Provider A across various socioeconomic patient strata, especially the most socioeconomically vulnerable stratum.

Appendix C – Measure Evaluation Summary Tables

Measures Recommended

0171 Acute Care Hospitalization During the First 60 Days of Home Health

Submission | Specifications

Description: Percentage of home health stays in which patients were admitted to an acute care hospital during the 60 days following the start of the home health stay.

Numerator Statement: Number of home health stays for patients who have a Medicare claim for an unplanned admission to an acute care hospital in the 60 days following the start of the home health stay.

Denominator Statement: Number of home health stays that begin during the 12-month observation period.

Exclusions: The following are excluded:

1) Home health stays for patients who are not continuously enrolled in fee-for-service Medicare for the 60 days following the start of the home health stay or until death.

2) Home health stays that begin with a Low Utilization Payment Adjustment (LUPA) claim.

3) Home health stays in which the patient receives service from multiple agencies during the first 60 days.

4) Home health stays for patients who are not continuously enrolled in fee-for-service Medicare for the 6 months prior to the home health stay.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Multinomial logit with outcomes of "No acute event", "Emergency Department without Hospitalization", and "Acute Care Hospitalization".

Level of Analysis: Facility

Setting of Care: Home Health

Type of Measure: Outcome

Data Source: Administrative claims

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

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

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

1a. Evidence: **Y-18; N-0**; 1b. Performance Gap: **H-3**; **M-15**; **L-0**; **I-0**

Rationale:

- The developer provided data on the distribution of performance of this measure for four years (2011, 2012, 2013, and 2014). These data note that the average risk-adjusted acute care hospitalizations for 2014 were 14.8%; and the 25th percentile was 12.7% and 75th percentile was 16.8%. This distribution of agency performance has a standard deviation of 3.3%. Based on these results the Standing Committee concurred a gap in care exists and that there is an opportunity for improvement.
- The Standing Committee noted that there is evidence that home health agencies can implement interventions to reduce admissions and that a performance gap exists. The Standing Committee also noted that performance on the measure varies across facilities.

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-2**; **M-16**; **L-0**; **I-0** 2b. Validity: **H-1**; **M-17**; **L-1**; **I-0**

Rationale:

0171 Acute Care Hospitalization During the First 60 Days of Home Health

- A beta-binomial distribution was fitted for all agencies. The beta-binomial method was developed for provider level measures reported as rates, and it allows one to calculate an agency level "reliability score," interpreted as the percent of variance due to the difference in measure score among providers.
    - The developer notes that the distribution of national reliability scores shows that the majority of agencies have a reliability score greater than 0.871 and that this implies their performance can likely be distinguished from other agencies. This can be interpreted as 87% of the variance is due to differences among providers, and 13% of the variance is due to measurement error or sampling uncertainty.
- The validity of this measure was calculated at the measure score level using empirical testing. The developer did not conduct additional validity testing of the measure elements noting that CMS audits a sample of claims for acute inpatient hospitalizations as a part of the annual payment error calculations.
    - The developers tested the validity of the measure through the use of payment error audits. The developers justified this during the prior review by stating that there is no reason to believe hospital would be more likely to have erroneous claims for home health patients than for others.
- This measure employs a multinomial logit model for risk adjustment. Variables included in the model include prior care setting (e.g., outpatient emergency room, inpatient acute, psychiatric facility, etc.), health status (measured using HCCs and all remaining CCs), demographic information (measured using age-gender interactions), enrollment status (ESRD and disability), and interactions between these factors. The c-statistic is 0.693.
- The developer submitted a conceptual rationale for SDS adjustment but ultimately chose not to include SDS factors in the risk adjustment model based on limited impact on performance rates.
- The Standing Committee raised concerns that the availability of home health services and the question of which patients are accepted into home health could impact the validity of this measure. The Standing Committee noted that home health agencies have more flexibility about whether or not to accept a patient than other providers may have. However, the Standing Committee noted that in some markets hospitals are working with home health agencies to improve care coordination and to assist them in handling more complex patients.
- The Standing Committee agreed this measure met the reliability and validity criteria.

3. Feasibility: H-16; M-3; L-0; 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:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure is feasible to collect and implement.

4. Usability and Use: H-3; M-16; L-0; I-0

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

Rationale:

- This measure is currently publicly reported and is used in an accountability program. The measure is currently used for quality improvement and benchmarking.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-19**; **N-0**

Rationale

- The Standing Committee recognized the importance of reducing the number of home health stays for patients who have a Medicare claim for an unplanned admission to an acute care hospital in the 60 days following the start of the home health stay and recommended the measure for endorsement.

6. Public and Member Comment

- This measure received three comments. Two commenters expressed their agreement with the

0171 Acute Care Hospitalization During the First 60 Days of Home Health

endorsement of this measure. One commenter raised concerns about the measure developer's decision not to include socioeconomic factors in the risk adjustment model. The commenter also raised concerns about availability of home health services and the flexibility of home health agencies to choose whether not to accept a patient.

- Committee Response:
  - The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.
  - The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.
  - The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-X; N-X

9. Appeals

0173 Emergency Department Use without Hospitalization During the First 60 Days of Home Health

Submission | Specifications

Description: Percentage of home health stays in which patients used the emergency department but were not admitted to the hospital during the 60 days following the start of the home health stay.

Numerator Statement: Number of home health stays for patients who have a Medicare claim for outpatient emergency department use and no claims for acute care hospitalization in the 60 days following the start of the home health stay.

Denominator Statement: Number of home health stays that begin during the 12-month observation period.

Exclusions: The following are excluded:

1) Home health stays for patients who are not continuously enrolled in fee-for-service Medicare for the 60 days following the start of the home health stay or until death.

2) Home health stays that begin with a Low Utilization Payment Adjustment (LUPA) claim.

3) Home health stays in which the patient receives service from multiple agencies during the first 60 days.

4) Home health stays for patients who are not continuously enrolled in fee-for-service Medicare for the 6 months prior to the home health stay.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Multinomial logit with outcomes of "No acute event", "Emergency Department use but no Hospitalization", and "Acute Care Hospitalization".

0173 Emergency Department Use without Hospitalization During the First 60 Days of Home Health

Level of Analysis: Facility

Setting of Care: Home Health

Type of Measure: Outcome

Data Source: Administrative claims

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

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

<u>Rationale</u>:

- The developer provided data on the distribution of risk-adjusted performance on this measure for 2011-2014. The average risk-adjusted performance is 11.9%, with the 25th percentile performance at 11.1% and the 75th performance at 12.5%. Based on these results the Standing Committee concurred a gap in care exists and that there is an opportunity for improvement.
- Standing Committee members expressed concerns that there was a limited evidence base for this measure. Standing Committee members noted challenges patients encounter connecting with primary care providers and the limited demonstrated impact of interventions such as medication reconciliation, education, and falls prevention. However, the Standing Committee felt this is an important tracking measure that can provide important information about patients' ability to provide self-care to remain stable in the community setting.
- The Standing Committee noted that tracking ED use could become an important issue as the healthcare system moves to alternative payment models. The Standing Committee also noted that not all referrals to the ED should be seen as a bad thing as this can represent the home health agency recognizing an acute problem early and getting the patient to the appropriate level of care.
- The Standing Committee noted that results on this measure are not improving but that this could be due to the patient population getting sicker over time. The Standing Committee suggested the developer better track data for multiple chronic conditions and co-morbidities.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-0**; **M-17**; **L-0**; **I-0** 2b. Validity: **H-1**; **M-16**; **L-0**; **I-0**

<u>Rationale</u>:

- A beta-binomial distribution was fitted for all agencies. The beta-binomial method was developed for provider level measures reported as rates, and it allows one to calculate an agency level "reliability score," interpreted as the percent of variance due to the difference in measure score among providers.
- The developer noted that the distribution of national reliability scores shows that the majority of agencies have a reliability score greater than 0.818 and that this implies their performance can likely be distinguished from other agencies.
    - This can be interpreted as approximately 82% of the variance is due to differences among providers, and 12% of the variance is due to measurement error or sampling uncertainty.
- The developer performed an audit of claims data to test the validity of the measure score. Of a 2010 audit of 31,766 Part B claims, there was 0.2% (801) claims that can patient record could not be found.
- This measure employs a multinomial logit model for risk adjustment. Variables included in the model include prior care setting (e.g., outpatient emergency room, inpatient acute, psychiatric facility, etc.), health status (measured using HCCs and all remaining CCs), demographic information (measured using age-gender interactions), enrollment status (ESRD and disability), and interactions between these factors. The c-statistic is 0.632.
- The developer submitted a conceptual rationale for SDS adjustment but ultimately chose not to include SDS factors in the risk adjustment model based on limited impact on performance rates.
- The Standing Committee suggested the developer look to other sources of data such as the Continuity of

0173 Emergency Department Use without Hospitalization During the First 60 Days of Home Health
Care Document to improve the risk models for this measure. • The Standing Committee agreed this measure met the reliability and validity criteria.
3. Feasibility: H-14; M-2; L-0; **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: • This measure is collected through administrative claims data. The Standing Committee agreed the measure is feasible to collect and implement.
4. Usability and Use: H-7; M-9; L-0; **I-0** *(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)* Rationale: • This measure is currently publicly reported and is used in an accountability program. The measure is currently used for quality improvement and benchmarking.
5. Related and Competing Measures • The Standing Committee raised concerns that this measure may compete with NQF #2505.The developer stated that this measure is "harmonized with the Rehospitalization measures (NQF numbers 2505 and 2380) and with CMS' Hospital-Wide All-Cause Unplanned Readmission (HWR) measure (NQF 1789) in the definition of unplanned hospitalizations." The developer added that this measure differs from other post-acute hospital readmission measures due to the unique nature of home health care as a post-acute setting.
Standing Committee Recommendation for Endorsement: **Y-16**; **N-0** Rationale • The Standing Committee recognized the importance of reducing the number of avoidable emergency department visits for the elderly without readmission among the elderly community and recommended this measure for continued endorsement.
6. Public and Member Comment • This measure received one comment in support of its endorsement.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization
Submission | Specifications
Description: The measure estimates a hospital-level risk-standardized readmission rate (RSRR) for patients discharged from the hospital with a principal diagnosis of heart failure (HF). The outcome (readmission) is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission (the admission included in the measure cohort). A specified set of planned readmissions do not count in the readmission outcome. The target population is patients 18 and over. CMS annually reports the measure for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare, and hospitalized in non-federal hospitals or Veterans Health Administration (VA) hospitals. **Numerator Statement**: The outcome for this measure is 30-day readmission. We define readmission as an inpatient admission for any cause, with the exception of certain planned readmissions, within 30 days from the date of discharge from the index HF admission. If a patient has more than one unplanned admissions (for any reason) within 30 days after discharge from the index admission, only one is counted as a readmission. The measure looks for a dichotomous yes or no outcome of whether each admitted patient has an unplanned

0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization

readmission within 30 days. However, if the first readmission after discharge is considered planned, any subsequent unplanned readmission is not counted as an outcome for that index admission, because the unplanned readmission could be related to care provided during the intervening planned readmission rather than during the index admission.

Denominator Statement: This claims-based measure can be used in either of two patient cohorts: (1) patients aged 65 years or older or (2) patients aged 18 years or older. We have explicitly tested the measure in both age groups.

The cohort includes admissions for patients aged 18 years and older discharged from the hospital with either a principal discharge diagnosis of HF (see codes below) and with a complete claims history for the 12 months prior to admission. The measure is currently publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals or Veterans Health Administration (VA) hospitals.

Additional details are provided in S.9 Denominator Details.

Exclusions: The readmission measures excludes admissions:

1. Ending in discharges against medical advice

Rationale: Providers did not have the opportunity to deliver full care and prepare the patient for discharge.

2. Without at least 30 days of post-discharge enrollment in FFS Medicare

Rationale: The 30-day readmission outcome cannot be assessed in this group since claims data are used to determine whether a patient was readmitted.

3. Occurring within 30 days of discharge from an index admission

Rationale: This exclusion ensures that no hospitalization will be considered as both a readmission and an index admission within the same measure.

4. With a procedure code for LVAD implantation or heart transplantation either during the index admission or in the 12 months prior to the index admission

Rationale: Patients with these procedures are a highly-selected group of patients with a different risk of the readmission outcome.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-19; N-1** 1b. Performance Gap: **H-9**; **M-9**; **L-1**; **I-1**

<u>Rationale</u>:

- Data provided by the developer cover a total of 1,210,454 and show that heart failure readmission rates ranges from a minimum of 16% to a maximum of 32.1%.
- Hospitals serving low proportion of Dual Eligible, African-American, and patients below AHRQ SES index score of 42.7 had lower readmission rates than those with high proportions of these patients.
- The Standing Committee discussed the two updates to the measure. First, the updated measure excludes patients who have either an LVAD or a heart transplant during their indexed stay or during the year prior. The Standing Committee generally agreed that this change was an appropriate reflection of a change in clinical practice. Second, the measure had modest changes to the planned readmissions algorithm which

0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization

excludes scheduled or planned readmissions from the measure.

- The noted that there is still a performance gap, with the average heart failure readmission rate over 22 percent and rates ranging from 16 percent to over 32 percent.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-7**; **M-12**; **L-0**; **I-1** 2b. Validity: **H-1**; **M-17**; **L-1**; **I-1**

<u>Rationale</u>:

- With regard to data element reliability, the developer notes that the measure has been developed to avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS.
- In addition, the developer compared frequencies and odds ratios of variables from their risk model across three years of data in order to assess the consistency of those variables over time.
- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method. This is generally considered an appropriate method of testing reliability.
- A total of 1,210,454 admissions over a 3-year period were examined, with 604,022 in one sample and 606,432 in the other randomly-selected sample. Two risk-standardized readmission rates (RSRR) were calculated for each hospital: one from each of the two separate samples. The agreement between the two RSRRs for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.58.
- The developer demonstrated measure validity through medical record validation.
    - The HF readmission administrative model (original model specification prior to completion of the planned readmission algorithm) was validated against a medical record model with the same cohort of patients for whom hospital-level HF readmission medical record data are available.
    - A measure cohort was developed with medical record data using the inclusion/exclusion criteria and risk-adjustment strategy.
    - A sample of 64,329 patients was matched for comparison.
- This measure employs a hierarchical logistic regression model (a form of hierarchical generalized linear model [HGLM]) to create a hospital level 30-day risk-standardized readmission rate (RSRR). Variables considered for inclusion in the model were patient-level risk-adjustors that are expected to be predictive of readmission, based on empirical analysis, prior literature, and clinical judgment, including demographic factors (age, sex) and indicators of comorbidity and disease severity. The C-statistic is 0.63.
- The developers also considered a number of variables related to sociodemographic status (SDS) for potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources.
- SDS variables were ultimately not included in model as the developer found that the effect size of each of these variables is small, the c-statistic (i.e., predictive value) is unchanged with the addition of any of these variables into the model, and the addition of any of these variables into the model has little to no effect on hospital performance.
- The Standing Committee expressed concerns about published literature suggesting there was a small, but significant inverse correlation between readmissions and mortality and recommended continued monitoring.
- Overall, the Standing Committee agreed this measure met the reliability and validity criteria.

3. Feasibility: H-16; M-4; L-;0 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)

<u>Rationale</u>:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure would be feasible to collect and implement.

0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization

4. Usability and Use: H-3; M-15; L-2; I-0

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

Rationale:

- It is currently used in the Hospital Inpatient Quality Report (IQR) and Hospital Readmissions Reduction (HRRP) Programs.
- The Standing Committee noted that this measure is associated with reduction in hospital RSRR by 1.6% between 2011-2012 and 2013-2014.
- The Standing Committee agreed the measure is highly usable.

5. Related and Competing Measures

- This measure is related to #2880: Excess days in acute care (EDAC) after hospitalization for heart failure.

Standing Committee Recommendation for Endorsement: **Y-19**; **N-1**

Rationale

- The Standing Committee recognized the importance of reducing readmissions due to heart failure and the need for improved care coordination and recommended the measure for continued endorsement.

6. Public and Member Comment

- This measure received four comments. One commenter raised concerns about the inverse correlation between readmissions and mortality for heart failure. The commenter also raised concerns that the Intra-Class Correlation Coefficient (ICC) for this measure was .58. The commenter also questioned that the C-statistic for this measure was 0.63.
- Three commenters noted the need for this measure to be adjusted for SDS factors.
- One commenter raised concerns that this measure could be implemented at levels of analysis other than the one for which it is endorsed. This commenter also raised concerns that this measure competes with #0277.
- One commenter raised concerns about the relationship between declining hospital admission rates and readmissions.
- Developer response:
    - Inverse correlation between readmissions and mortality
        - The hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) and risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization have been publicly reported since June 2007 and June 2009, respectively. Yale-CORE reported the results of an examination of the correlation between the two outcomes using CMS claims data from 2005-2008 in a published study (Krumholz HM, Lin Z, Keenan PS, et al. Relationship between hospital readmission and mortality rates for patients hospitalized with acute myocardial infarction, heart failure, or pneumonia. JAMA 2013; 309:587-593). The results demonstrated that the correlation, although statistically significant, is relatively low (the Pearson correlation is -0.17 with a 95% CI of -0.20 to -0.14) and only exists in the lower range of RSMRs. The much more dominant finding is that hospitals can perform well on both measures and that a relatively important share of hospitals perform above the national average or below the national average on both mortality and readmission measures. These results, which are consistent across different types of hospitals, such as teaching hospitals and rural hospitals, demonstrate that there is no systematic relationship between the two measures.
    - Intra-Class Correlation Coefficient
        - We used the Inter-Class Correlation (ICC) method to establish the reliability of the measure score. Our approach to assessing reliability is to consider the extent to which assessments of a hospital using different but randomly selected subsets of patients produces similar measures of hospital performance. That is, we take a "test-retest" approach in which hospital performance is measured once using a random subset of

patients, then measured again using a second random subset exclusive of the first, and finally comparing the agreement between the two resulting performance measures across hospitals (Rousson V, Gasser T, Seifert B. Assessing intrarater, interrater and test–retest reliability of continuous measurements. Statistics in Medicine 2002;21:3431-3446.). This is a purposefully conservative approach to assessing reliability and traditional thresholds for acceptability do not apply to interpreting these results.
        - The minimally acceptable threshold noted by AHS is not appropriate for this particular analytic approach. We have cited the more appropriate convention, which describes the ICC values as moderate (0.41-0.60) for this measure (Landis JR and Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977; 33:159-174).
    - SDS adjustment
        - CMS and Yale-CORE examined heart failure readmission measure results, or hospitals' performance on this measure, using their entire patient populations including both patients with and without low SES risk variables and we found observed that hospitals had similar performance in both groups. Additionally, we examined the impact of adding patient-level risk adjustment which aims to answer the extent to which patients' SES affects measure results and found very little difference in hospital scores. We also examined risk models that included all patient comorbid conditions, both SES variables (dual eligibility and AHRQ SES Index Score) and African-American or non-African-American race, and found no change to the c-statistics compared with models that did not include SES and race variables.
    - C-statistic
        - A higher C-statistic is not always better for outcome quality measures. The goal of the measures is to assess quality by estimating hospital outcome rates and accounting for important patient factors. It is not to produce the best model for predicting patient outcomes. Considering an extreme example of an outcome which is fully determined by hospital care and not at all influenced by patient risk factors of any sort, we would in that case expect to observe a C-statistic of 0.5. But if hospital quality, not patient factors, were responsible for the outcome we would still conclude that this was a good quality measure, but simply that risk adjustment was unnecessary. In the case of the readmission measures, patient factors are not particularly strong predictors of the readmission outcome and our C-statistics for readmission are consistent with those reported in the literature as appropriate for assessing quality (Kansagara D., Englander H., Salatrino A., et al. Risk Prediction Models for Hospital Readmission A Systematic Review. JAMA 2011; 306(15): 1688-1698; Bradley E, Yakusheva O, Horwitz LI, Sipsma H, Fletcher J. Identifying Patients at Increased Risk for Unplanned Readmission. Medical care. 2013;51(9):761-766). A crucial additional note is that, because differences in RSRRs are intended to reflect differences in quality of care among hospitals, we purposefully do not account for any aspect of the care patients receive in our risk models. You can achieve a higher C-statistic by adding information about care received to the risk model, such as interventions but such models would provide less ability to illuminate differences in quality across hospitals as they would adjust away some of the quality signal. Similarly, we could increase C-statistics by including in-hospital complications as risk factors, but it would be inappropriate for the purpose of assessing hospital quality of care.
- Committee response:
    - The Committee has reviewed your comment and appreciates your input. Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as

0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization

developed by the measure developer.

The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

- The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case.
- The Committee followed NQF's guidance on measure harmonization throughout the evaluation process. Prior to the in-person meeting, the Committee received materials regarding these competing measures, and held a separate call after the in-person meeting on September 1 to discuss harmonization issues and allow the developers to answer questions from Committee members. The Committee then voted via survey to recommend both measures. The Committee considered the added value and burden of recommending both measures and agreed that the differences in measure specifications added sufficient value to offset any potential negative impact.
- The Committee has reviewed your comment and appreciates your input. The Committee recognizes the potential for negative unintended consequences of admissions and readmissions measures and recommends careful monitoring of their implementation. Above all, the Committee agreed that use of these measures should be monitored to ensure they do not inadvertently reduce access to necessary care. The Committee noted the inverse relationship between mortality and readmission for heart failure and recognized the need for careful surveillance and balancing of these measures. The Committee also reiterated its concerns about the need to carefully balance implementation of measures addressing psychiatric readmissions to prevent the risk of higher suicidality.

  On the other hand, the Committee has noted the desire to understand a patient's need for any subsequent acute care after a hospitalization. In particular, the Committee recognized the need understand if patients are being seen in the Emergency Department after discharge or being

0330 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following heart failure (HF) hospitalization
placed in observation. The Committee recommends continued work to ensure that the use of readmissions measures does not result in unnecessary or avoidable use of the emergency department or observation status while ensuring that all patients have access to any necessary care. The Committee noted that a number of measures recommended for endorsement in this project could help to balance these concerns, in particular the measures addressing excess days in acute care and population-based admission measures.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

0506 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization
Submission | Specifications
Description: The measure estimates a hospital-level 30-day, all-cause, risk-standardized readmission rate (RSRR) for patients discharged from the hospital with either a principal discharge diagnosis of pneumonia, including aspiration pneumonia or a principal discharge diagnosis of sepsis (not severe sepsis) with a secondary diagnosis of pneumonia (including aspiration pneumonia) coded as present on admission (POA). Readmission is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission. A specified set of planned readmissions do not count as readmissions. CMS annually reports the measure for patients who are 65 years or older and are enrolled in fee-for-service (FFS) Medicare hospitalized in non-federal hospitals. Please note this measure has been substantially updated since the last submission; as described in S.3., the cohort has been expanded. Throughout this application we refer to this measure as version 8.2. **Numerator Statement**: The outcome for this measure is 30-day readmission. We define readmission as an inpatient admission for any cause, with the exception of certain planned readmissions, within 30 days from the date of discharge from the index admission for patients 18 and older discharged from the hospital with a principal discharge diagnosis of pneumonia, including aspiration pneumonia or a principal discharge diagnosis of sepsis (not severe sepsis) with a secondary discharge diagnosis of pneumonia (including aspiration pneumonia) coded as POA and no secondary discharge diagnosis of severe sepsis. If a patient has more than one unplanned admission (for any reason) within 30 days after discharge from the index admission, only the first one is counted as a readmission. The measure looks for a dichotomous yes or no outcome of whether each admitted patient has an unplanned readmission within 30 days. However, if the first readmission after discharge is considered planned, any subsequent unplanned readmission is not counted as an outcome for that index admission because the unplanned readmission could be related to care provided during the intervening planned readmission rather than during the index admission. **Denominator Statement**: This claims-based measure can be used in either of two patient cohorts: (1) patients aged 65 years or over or (2) patients aged 18 years or older. We have specifically tested the measure in both age groups. The cohort includes admissions for patients aged 18 years and older discharged from the hospital with principal discharge diagnosis of pneumonia, including aspiration pneumonia or a principal discharge diagnosis of sepsis (not severe sepsis) with a secondary discharge diagnosis of pneumonia (including aspiration pneumonia) coded as POA and no secondary discharge diagnosis of severe sepsis; and with a complete claims history for the 12 months prior to admission. The measure will be publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals. Additional details are provided in S.9 Denominator Details. **Exclusions**: The readmission measures exclude index admissions for patients: 1. Discharged against medical advice (AMA); 2. Without at least 30 days post-discharge enrollment in FFS Medicare; 3. Admitted within 30 days of a prior index admission. **Adjustment/Stratification**: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the

0506 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization

American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

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

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

1a. Evidence: **Y-20; N-0**; 1b. Performance Gap: **H-12**; **M-7**; **L-1**; **I-0**

Rationale:

- New evidence is provided since the last endorsement maintenance review. Since its last review, this measure has been updated to include an expanded cohort to include patients with aspiration pneumonia and sepsis.
- Data provided by the developer cover a total of 1,469,277 and show that pneumonia readmission rates ranges from a minimum of 13.1% to a maximum of 24.7%.
- Hospitals serving low proportion of Dual Eligible, African-American, and patients below AHRQ SES index score of 42.7 had lower readmission rates than those with high proportions of these patients.
- The Standing Committee reviewed the two measure updates. First, the measure has an expanded cohort including patients who have a principal diagnosis of sepsis and a secondary diagnosis of pneumonia that is present on admission, and patients who have a principal diagnosis of aspiration pneumoia. Second, the measure includes the updated planned readmissions algorithm noted for Measure #0330.
- The Standing Committee agreed that the measure still has a performance gap, with rates of pneumonia readmission ranging from 13.1 percent to 24.7 percent with an average rate of 17.5 percent.

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-12**; **M-8**; **L-0**; **I-0** 2b. Validity: **H-1**; **M-17**; **L-1**; **I-0**

Rationale:

- With regard to data element reliability, the developer notes that the measure has been developed to avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS.
- In addition, the developer compared frequencies and odds ratios of variables from their risk model across three years of data in order to assess the consistency of those variables over time.
- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method. This is generally considered an appropriate method of testing reliability.
- A total of 1,469,277 admissions over a 3-year period were examined, with 733,434 in one sample and 735,843 in the other randomly-selected sample. Two risk-standardized mortality rates (RSMR) were calculated for each hospital: one from each of the two separate samples. The agreement between the two RSMRs for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.73.
- The developer tested the original version of the measure by comparing the administrative model with a medical-record based model. The results of this testing are included in the citation Krumholz, 2008. The developer notes that the claims-based measure produced results which were highly correlated with those produced through manual chart audit. (Krumholz et al., 2008; Lindenauer et al., 2011)
- This measure employs a hierarchical logistic regression model (a form of hierarchical generalized linear model [HGLM]) to create a hospital level 30-day risk-standardized readmission rate (RSRR). Variables considered for inclusion in the model were patient-level risk-adjustors that are expected to be predictive of readmission, based on empirical analysis, prior literature, and clinical judgment, including demographic factors (age, sex) and

0506 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization

indicators of comorbidity and disease severity. The C-statistic is 0.63.

- The developers also considered a number of variables related to sociodemographic status (SDS) for potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources.
- SDS variables were ultimately not included in model as the developer found that the effect size of each of these variables is small, the c-statistic (i.e., predictive value) is unchanged with the addition of any of these variables into the model, and the addition of any of these variables into the model has little to no effect on hospital performance.
- The Standing Committee questioned whether hospitals with a larger proportion of aspiration pneumonia patients did similar to other hospitals, to which the developers noted yes. Additionally, Standing Committee members expressed concerns on the lack of data published on sensitivity and specificity for patients with sepsis and pneumonia as a secondary diagnosis in hospitals. Overall, the Standing Committee agreed this measure had sufficient reliability and validity testing to meet the criteria.

3. Feasibility: H-15; M-4; L-0; 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:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure would be feasible to collect and implement.

4. Usability and Use: H-4; M-15; L-0; I-0

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

Rationale:

- It is currently used in the Hospital Inpatient Quality Report (IQR) and Hospital Readmissions Reduction (HRRP) Programs.
- The Standing Committee agreed the measure is highly usable.

5. Related and Competing Measures

- This measure is related to NQF #0279: Bacterial Pneumonia Admission Rate (PQI 11) and NQF #2882: Excess days in acute care (EDAC) after hospitalization for pneumonia. The developer notes that the measures are not completely harmonized. The developer justifies the difference by noting that for outcome measures clinical coherence of the cohort takes precedence over alignment with related non-outcome measures.

Standing Committee Recommendation for Endorsement: **Y-18**; **N-1**

Rationale

- The Standing Committee recognized the importance of reducing readmissions due to pneumonia and the need for improved care coordination and discharge management and recommended the measure for continued endorsement.

6. Public and Member Comment

- Four comments were received on this measure. One commenter raised concerns about the expansion of the measure cohort. Two commenters expressed concerns that the measure does not include SDS factors in the risk adjustment model. One commenter noted that his measure should account for planned admissions. One commenter raised concerns about the relationship between declining admission rates and readmissions.
- Developer response:
  - Expanded cohort:
    - Several studies in the published literature have shown a rapid increase in the use of sepsis codes over recent years for patients with pneumonia, and wide variation in the use of sepsis as a principal discharge diagnosis code across hospitals (see, e.g., Sjoding MW, Iwashyna TJ, Dimick JB, et al. Gaming hospital-level pneumonia 30-day mortality and readmission measures by legitimate changes to diagnostic coding. Crit Care Med. 2015; 43(5): 989-995). Analyses conducted by Yale-CORE as a part of measure reevaluation demonstrated that expansion of

the cohorts to include patients with sepsis and aspiration pneumonia appeared to mitigate or resolve bias in hospital performance related to diagnostic coding patterns. For example, our analyses confirmed findings of previous studies showing that hospital coding for sepsis was strongly associated with hospital performance on the pneumonia mortality and readmission measures. These findings suggested that at some hospitals where the sepsis code was used more frequently, patients who met the diagnostic definition for sepsis and also had pneumonia were being excluded from the measure. This pattern of excluding potentially sicker pneumonia patients from the measure cohort was biasing the measure in favor of hospitals with high rates of sepsis coding. We found that this issue was resolved by the addition of the sepsis patients to the measure. Patients with pneumonia severe enough to be admitted to the hospital frequently meet criteria for sepsis and should be a part of the measure cohort. However, we do not include patients with severe sepsis.

- Similar patterns were found in aspiration pneumonia. Hospitals used aspiration pneumonia as a principal discharge diagnosis code to varying degrees and therefore not including these patients in the measure could lead to differential exclusions across hospitals. Additionally, there is no commonly accepted definition or gold standard diagnostic test to identify aspiration pneumonia. This is a subset of bacterial pneumonia which is diagnosed clinically, often subjectively based on patient's risk factors, such as age and frailty. The treatment of patients who receive a diagnosis for aspiration pneumonia is carried out by the same care teams and using similar approaches as patients with other types of pneumonia. Additionally, the prospective payment system creates strong incentives for hospitals to make a diagnosis of aspiration pneumonia because it changes the Diagnostic Related Group (DRG) from simple pneumonia (MS-DRG 177-179) to a higher reimbursement DRG for respiratory infections and inflammations (MS-DRG 193-195). Variation across hospitals in the application of the aspiration pneumonia code has the potential to bias outcomes estimated across hospitals when calculated for the mortality and readmission measures, in the same way that variation in sepsis coding has been shown to introduce bias in the pneumonia measures.
- Our findings argued for the need to broaden the cohort to capture the full spectrum of disease presentation and to reduce potential bias. The rationale for expansion was based on the intent, to include the full spectrum of patients admitted for the treatment of pneumonia and to prevent bias based on different coding patterns.

- Planned readmission
  - The CMS readmission measures do not consider planned readmissions as part of the readmission outcome. Generally speaking, planned readmissions are not a signal of quality of care. Therefore, CMS has worked with experts in the medical community as well as other stakeholders to carefully identify procedures and treatments that should be considered "planned," and thus not considered in the readmission outcome. Starting with the 2013 public reporting, the measures identify planned readmissions by using an expanded algorithm, which is a set of criteria for classifying readmissions as planned using Medicare claims. This algorithm identifies admissions that are typically planned and may occur within 30 days of discharge from the hospital. The algorithm is based on three principles: • A few specific, limited types of care are always considered planned (transplant surgery, maintenance chemotherapy/radiotherapy/ immunotherapy, rehabilitation);• Otherwise, a planned readmission is defined as a non-acute readmission for a scheduled procedure; and • Admissions for acute illness or for complications of care are never planned. CMS conducted a validation study of the planned readmission algorithm using medical record data from 634 medical records at seven hospitals. For the 2016 public reporting, Version 4.0 of the algorithm includes modifications to enhance the accuracy of the algorithm based on the study findings. These changes improve the accuracy of the algorithm by decreasing the number of readmissions that the algorithm mistakenly designates as planned or unplanned. This involved the removal of five procedure categories and the addition one procedure category to the list of potentially planned procedure that disqualify readmissions from the measure outcome.

        - For the details of the planned readmission algorithm as applied to the pneumonia measure, refer to Appendix E of the 2016 Condition-Specific Readmission Measures Updates and Specifications Report available on QualityNet at: (www.qualitynet.org) > Hospitals – Inpatient > Claims-Based Measures > Readmission Measures > 2016 AMI, HF, Pneumonia, COPD, and Stroke Readmission Measures Updates and Specifications Report (also available at http://www.qualitynet.org/dcs/BlobServer?blobkey=id&blobnocache=true&blobwhere=1228890567694&blobheader=multipart%2Foctet-stream&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3Bfilename%3DCondSpecific_Rdmsn_Rpt_2016.pdf&blobcol=urldata&blobtable=MungoBlobs).
    - SDS adjustment
        - CMS agrees that patients' socioeconomic status (SES) affects health and health outcomes in important ways. In the conceptual model presented to the Committee, we explain that many patients with low SES indicators may have poorer health status at the start of an index admission that increases their risk of readmission. The decrease in the strength of the association between SES variables and the readmission outcome when we added patients' comorbidities to the risk model supports this proposed mechanism. Additionally, the results presented showed that the effect of SES variables on readmission rates in the multi-variate or fully adjusted model was small but significant. However, inclusion of these variables did not change hospitals risk-standardized readmission rates or their performance on the measures. Yale-CORE remains committed to examining alternative solutions that better reflect the balance of hospital- and patient-level influences on hospital outcome measures and to considering appropriate ways to incorporate community factors into the outcomes measures.
- Committee response:
    - Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging. The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

      The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.
    - The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.The Committee recognizes the potential for negative unintended consequences of admissions and readmissions measures and recommends careful monitoring of their implementation.

0506 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization
Above all, the Committee agreed that use of these measures should be monitored to ensure they do not inadvertently reduce access to necessary care. The Committee noted the inverse relationship between mortality and readmission for heart failure and recognized the need for careful surveillance and balancing of these measures. The Committee also reiterated its concerns about the need to carefully balance implementation of measures addressing psychiatric readmissions to prevent the risk of higher suicidality. On the other hand, the Committee has noted the desire to understand a patient's need for any subsequent acute care after a hospitalization. In particular, the Committee recognized the need understand if patients are being seen in the Emergency Department after discharge or being placed in observation. The Committee recommends continued work to ensure that the use of readmissions measures does not result in unnecessary or avoidable use of the emergency department or observation status while ensuring that all patients have access to any necessary care. The Committee noted that a number of measures recommended for endorsement in this project could help to balance these concerns, in particular the measures addressing excess days in acute care and population-based admission measures.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)
Submission | Specifications
Description: The measure estimates a hospital-level risk-standardized readmission rate (RSRR) of unplanned, all-cause readmission after admission for any eligible condition within 30 days of hospital discharge. The measure reports a single summary RSRR, 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. The measure also indicates the hospital-level standardized risk ratios (SRR) for each of these five specialty cohorts. The outcome is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission (the admission included in the measure cohort). A specified set of planned readmissions do not count in the readmission outcome. CMS annually reports the measure for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare, and hospitalized in non-federal hospitals. **Numerator Statement**: The outcome for this measure is 30-day readmission. We define readmission as an inpatient admission for any cause, with the exception of certain planned readmissions, within 30 days from the date of discharge from an eligible index admission. If a patient has more than one unplanned admission (for any reason) within 30 days after discharge from the index admission, only one is counted as a readmission. The measure looks for a dichotomous yes or no outcome of whether each admitted patient has an unplanned readmission within 30 days. However, if the first readmission after discharge is considered planned, any subsequent unplanned readmission is not counted as an outcome for that index admission because the unplanned readmission could be related to care provided during the intervening planned readmission rather than during the index admission. **Denominator Statement**: The measure includes admissions for Medicare beneficiaries who are 65 years and older and are discharged from all non-federal, acute care inpatient US hospitals (including territories) with a complete claims history for the 12 months prior to admission. Additional details are provided in S.9 Denominator Details. **Exclusions**: The measure excludes index admissions for patients: 1. Admitted to Prospective Payment System (PPS)-exempt cancer hospitals; 2. Without at least 30 days post-discharge enrollment in FFS Medicare; 3. Discharged against medical advice (AMA); 4. Admitted for primary psychiatric diagnoses; 5. Admitted for rehabilitation; or

1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)

6. Admitted for medical treatment of cancer.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

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

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

1a. Evidence: **Unchanged – no vote** 1b. Performance Gap: **H-3**; **M-15**; **L-1**; **I-0**

Rationale:

- The developer stated that there are no updates to the evidence since the last submission, so the Standing Committee agreed that there was no need for a repeat discussion or vote on evidence.
- Data provided by the developer cover a total of 22,000,000 admissions and show that readmission rates ranges from a minimum of 11.4% to a maximum of 20.1%.
- Hospitals serving low proportion of Dual Eligible, African-American, and patients below AHRQ SES index score of 45 had lower readmission rates than those with high proportions of these patients.
- The Standing Committee agreed that there continues to be a performance gap, with all cause readmission rates ranging from 11.4 percent to 20.1 percent with an average of 15.4 percent.

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-8**; **M-11**; **L-0**; **I-0** 2b. Validity: **H-8**; **M-11**; **L-0**; **I-0**

Rationale:

- With regard to data element reliability, the developer notes that the measure has been developed to avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS.
- In addition, the developer compared frequencies and odds ratios of variables from their risk model across three years of data in order to assess the consistency of those variables over time.
- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method. This is generally considered an appropriate method of testing reliability.
- A total of 6,843,808 admissions in the 2015 publicly reported measure, with 3,420,728 in one sample and 3,423,080 in the other randomly-selected sample. Two risk-standardized readmission rates (RSRR) were calculated for each hospital: one from each of the two separate samples. The agreement between the two RSRRs for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.80.
- The developer demonstrated measure validity through prior validity testing done on their other claims-based measures, through the use of established measure development guidelines, and examination of content validity by comparing hospital performance with that on other quality measures.
- This measure employs a hierarchical logistic regression model (a form of hierarchical generalized linear model [HGLM]) to create a hospital level 30-day risk-standardized readmission rate (RSRR). Variables considered for inclusion in the model were patient-level risk-adjustors that are expected to be predictive of readmission, based on empirical analysis, prior literature, and clinical judgment, including demographic factors (age, sex) and indicators of comorbidity and disease severity.
- C-statistic for each cohort:

1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)

  - Medicine cohort: 0.643
  - Surgical cohort: 0.675
  - Cardiorespiratory cohort: 0.636
  - Cardiovascular cohort: 0.658
  - Neurology cohort: 0.622
- The developers also considered a number of variables related to sociodemographic status (SDS) for potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources.
- SDS variables were ultimately not included in model as the developer found that the effect size of each of these variables is small, the c-statistic (i.e., predictive value) is unchanged with the addition of any of these variables into the model, and the addition of any of these variables into the model has little to no effect on hospital performance.
- The Standing Committee raised concerns that merging multiple cohorts into one group may mask the individual variance properties of the individual cohorts.
- The Standing Committee expressed that the modeling was laid out very explicitly and well-specified and generally agreed that the measure had sufficient reliability and validity testing to meet the reliability and validity criteria.

3. Feasibility: H-17; M-2; L-0; 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:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure would be feasible to collect and implement.

4. Usability and Use: H-7; M-12; L-0; I-0

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

Rationale:

- It is currently used in the Hospital Inpatient Quality Report (IQR) Program.
- The Standing Committee agreed the measure is highly usable.

5. Related and Competing Measures

- This measure is related to NQF # 1768: Plan All-Cause Readmissions (PCR). This measure and the NCQA Plan All-Cause Readmissions (PCR) Measure #1768 are related measures, but are not competing because they don't have the same measure focus and same target population. Each of these measures has different specifications. In addition, both have been previously harmonized to the extent possible under the guidance of the National Quality Forum Standing Committee in 2011.

Standing Committee Recommendation for Endorsement: **Y-18**; **N-1**

Rationale

- The Standing Committee recognized the importance of reducing readmissions and the need for improved care coordination and discharge management and recommended the measure for continued endorsement.

6. Public and Member Comment

- Eight comments were received on this measure, including a sign-on letter from 28 physician societies. Two commenters expressed their support for endorsement of this measure.
- A number of commenters raised concerns that NQF #1789 Hospital-wide all-cause unplanned readmission measure is being used at the clinician level of analysis in the Physician Value-Based Payment Modifier program and is proposed to be used in the Merit-Based Incentive Payment System in a similar way. These commenters expressed concern that testing at this level of analysis was not provided to the Standing Committee for review.
- Two commenters raised concerns that this measure does not include SDS factors in the risk adjustment model.

1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)

- One commenter expressed concerns that trauma is not excluded from the measure.
- Developer Response:
  - SDS adjustment
    - CMS agrees that patients' socioeconomic status (SES) effects health and health outcomes in important ways. In the conceptual model presented to the Committee, we explain that many patients with low SES indicators may have poorer health status at the start of an index admission that increases their risk of readmission. The decrease in the strength of the association between SES variables and the readmission outcome when we added patients' comorbidities to the risk model supports this proposed mechanism. Additionally, the results presented showed that the effect of SES variables on readmission rates in the multi-variate or fully adjusted model was small but significant. However, inclusion of these variables did not change hospitals risk-standardized readmission rates or their performance on the measures. We explained that the remaining small effect of SES in the risk models could be a hospital-level effect, if patients with low SES indicators more often receive care at lower quality hospitals. Alternatively, it could be a patient-level effect, if patients have other unmeasured factors that increase their risk of readmission that are beyond the hospitals' control or if they receive inappropriate care from hospitals due to bias or discrimination. The results of the decomposition analyses we presented to the Committee confirmed that most of the small residual effect of SES variables on readmission rates is a hospital-level effect, suggesting that it is due to the clustering of patients with low SES indicators and low quality hospitals. Therefore, we concluded that the evidence did not support including SES variables in the measures risk models. We also note that the lack of any change in hospitals performance with inclusion of individual SES risk variables also held true when all SES variables were added to the fully adjusted model together. Yale-CORE remains committed to examining alternative solutions that better reflect the balance of hospital- and patient-level influences on hospital outcome measures and to considering appropriate ways to incorporate community factors into the outcomes measures.
  - Relationship between admission and readmission rates
    - In a recent study published in Health Affairs, Dharmarajan and colleagues (Dharmarajan K, Qin L, Lin ZQ, et al. Declining Admission Rates And Thirty-Day Readmission Rates Positively Associated Even Though Patients Grew Sicker Over Time. Health Affairs 2016; 35(7): 1294-1302) explore the relationship between admission and readmission rates. Using national data on Medicare fee-for-service beneficiaries from 2010 to 2013, the study shows that communities with a decline in admission rates also had a decline in readmission rates despite the fact that hospitalized patients were sicker. This association suggests that reducing admission rates does not necessarily lead to higher readmission rates. From a policy perspective, both outcomes might be pursued simultaneously.
  - 2015 Physician Fee Schedule Proposed Rule
    - Questions about application of this measure beyond the hospital setting is beyond the scope of what the developer was asked to examine and consider for measure endorsement maintenance.
  - Physician level reliability
    - This comment is out of scope for the hospital measures.
  - Trauma
    - The CMS readmission measures assess all-cause readmissions; that is, they consider unplanned readmissions for any reason, not only those that are due to the same or a "related" condition.
    - There are several reasons for measuring all-cause readmissions. First, from the patient perspective, an unplanned readmission is disruptive and costly regardless of cause. Second, restricting the measure outcomes to those readmissions that seem to be directly related to the initial hospitalization may make the measures susceptible to gaming through changes in coding practices. Although most hospitals would not engage in such practices, CMS wants to

eliminate any incentive for hospitals to change coding practices in an effort to prevent readmissions from being captured in their readmission measure results. Third, an apparently unrelated readmission may represent a complication related to the underlying condition. For example, a patient with heart failure who develops a hospital-acquired infection may later be readmitted due to that infection. It would be inappropriate to consider this readmission as unrelated to the care the patient received for heart failure. Finally, hospitals can act to reduce readmissions from all causes. While CMS does not presume that every readmission is preventable, measuring all-cause readmission incentivizes hospitals to evaluate the full range of factors that increase patients' risk for unplanned readmissions. For example, unclear discharge instructions, poor communication with post-acute care providers, and inadequate follow-up are factors that typically increase the risk for an unplanned readmission.

- Although measuring all-cause readmissions will include some patients whose readmission may be unrelated to their care (for example, a casualty in a motor vehicle accident), such events should occur randomly across hospitals and therefore will not affect results on measures that assess relative performance.
- Note that planned readmissions do not count as readmissions in the 30-day readmission measures. For the details of the planned readmission algorithm as applied to the HWR measures, refer to Appendix E of the 2016 Hospital-wide Readmission Measure Updates and Specifications Report available on QualityNet at: (www.qualitynet.org) > Hospitals – Inpatient > Claims-Based Measures > Readmission Measures > Measure Methodology > Hospital-Wide Readmission Measure Methodology Report (also available at http://www.qualitynet.org/dcs/BlobServer?blobkey=id&blobnocache=true&blobwhere=1228890434757&blobheader=multipart%2Foctet-stream&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3Bfilename%3DDryRun_HWR_TechReport_081012%2C0.pdf&blobcol=urldata&blobtable=MungoBlobs). Proposed Committee Response: Thank you for your comment. The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.

- Committee response:
  - The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.
  - The Committee recognizes the potential for negative unintended consequences of these measures and recommends careful monitoring of their implementation.
  - Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

    The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

    The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required

1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)
for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses. The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures. ○ The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

1891 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization
Submission | Specifications
Description: The measure estimates a hospital-level 30-day, all-cause, risk-standardized readmission rate (RSRR) for patients discharged from the hospital with either a principal discharge diagnosis of COPD or a principal discharge diagnosis of respiratory failure with a secondary diagnosis of acute exacerbation of COPD. The outcome (readmission) is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission (the admission included in the measure cohort). A specified set of planned readmissions do not count in the readmission outcome. CMS annually reports the measure for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare, and hospitalized in non-federal hospitals. **Numerator Statement**: The outcome for this measure is 30-day readmission. We define readmission as an inpatient admission for any cause, with the exception of certain planned readmissions, within 30 days from the date of discharge from the index admission for patients discharged from the hospital with a principal discharge diagnosis of COPD or principal discharge diagnosis of respiratory failure with a secondary discharge diagnosis of acute exacerbation of COPD. If a patient has more than one unplanned admission (for any reason) within 30 days after discharge from the index admission, only the first one is counted as a readmission. The measure looks for a dichotomous yes or no outcome of whether each admitted patient has an unplanned readmission within 30 days. However, if the first readmission after discharge is considered planned, any subsequent unplanned readmission is not counted as an outcome for that index admission because the unplanned readmission could be related to care provided during the intervening planned readmission rather than during the index admission. **Denominator Statement**: This claims-based measure can be used in either of two patient cohorts: (1) patients aged 65 years or older or (2) patients aged 40 years or older. We have explicitly tested the measure in both age groups. The cohort includes admissions for patients discharged from the hospital with either a principal discharge diagnosis of COPD (see codes below) OR a principal discharge diagnosis of respiratory failure (see codes below) with a secondary discharge diagnosis of acute exacerbation of COPD (see codes below) and with a complete claims

1891 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization

history for the 12 months prior to admission. The measure is currently publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals.

Additional details are provided in S.9 Denominator Details.

Exclusions: The readmission measures exclude index admissions for patients:

1. Without at least 30 days post-discharge enrollment in FFS Medicare.
2. Discharged against medical advice (AMA);
3. Admitted within 30 days of a prior index admission.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **unchanged – no vote**; 1b. Performance Gap: **H-7**; **M-13**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer states that there are no updates to the evidence since the last submission, so the Standing Committee agreed that there was no need for a repeat discussion on evidence.
- Data provided by the developer cover a total of 925,315 admissions and show that COP readmission rates ranges from a minimum of 15.5% to a maximum of 26.6%.
- Hospitals serving low proportion of Dual Eligible, African-American, and patients below AHRQ SES index score of 45 had lower readmission rates than those with high proportions of these patients.
- The Standing Committee agreed that the measure continues to have a performance gap with readmission rates for COPD ranging from 15.5 percent to 26.6 percent and an average of 20.2 percent.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-5**; **M-14**; **L-0**; **I-0** 2b. Validity: **H-8**; **M-12**; **L-0**; **I-0**

<u>Rationale</u>:

- With regard to data element reliability, the developer notes that the measure has been developed to avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS.
- In addition, the developer compared frequencies and odds ratios of variables from their risk model across three years of data in order to assess the consistency of those variables over time.
- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method. This is generally considered to be an appropriate method of testing reliability.
- A total of 925,315 admissions over a 3-year period were examined, with 461,505 in one sample and 463,810 in the other randomly-selected sample. Two risk-standardized readmission rates (RSRR) were calculated for each hospital: one from each of the two separate samples. The agreement between the

1891 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization

two RSRRs for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.48.

- The developer demonstrated measure validity through prior validity testing done on their claims-based measures, through use of established measure development guidelines, and by systematic assessment of measure face validity by a Technical Expert Panel (TEP).
- This measure employs a hierarchical logistic regression model (a form of hierarchical generalized linear model [HGLM]) to create a hospital level 30-day risk-standardized readmission rate (RSRR). Variables considered for inclusion in the model were patient-level risk-adjustors that are expected to be predictive of readmission, based on empirical analysis, prior literature, and clinical judgment, including demographic factors (age, sex) and indicators of comorbidity and disease severity. The C-statistic is 0.64.
- The developers also considered a number of variables related to sociodemographic status (SDS) for potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources.
- SDS variables were ultimately not included in model as the developer found that the effect size of each of these variables is small, the c-statistic (i.e., predictive value) is unchanged with the addition of any of these variables into the model, and the addition of any of these variables into the model has little to no effect on hospital performance.
- While there was discussion about the modest results of the reliability testing and the use of hierachical logistical modeling, the Standing Committee agreed that the measure met the reliability and validity criteria for NQF endorsement.

3. Feasibility: H-15 M-5 L-0; 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:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure would be feasible to collect and implement.

4. Usability and Use: H-9; M-11; L-0; I-0

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

Rationale:

- It is currently used in the Hospital Inpatient Quality Report (IQR) and Hospital Readmissions Reduction (HRRP) Programs.
- The Standing Committee noted that there are no unintended consequences for the measure, but had a few concerns regarding the advantages and disadvantages of the hierarchical approach how closely the predictive rate reflects hospital performance.
- Overall, the Standing Committee felt the measure met the NQF criteria for usability and use.

5. Related and Competing Measures

- This measure is related to NQF #0275: Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate (PQI 5).

Standing Committee Recommendation for Endorsement: **Y-18**; **N-2**

Rationale

- The Standing Committee recognized the importance of reducing readmissions due to COPD and the need for improved care coordination and discharge management and recommended the measure for continued endorsement.

6. Public and Member Comment

- Three comments were received on this measure. One commenter submitted a comment in support of recommending the measure for endorsement.
- One commenter raised concerns about the level of reliability for this measure, saying the Intra-Class

1891 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization

Correlation Coefficient (ICC) of 0.48 was low and an ICC of 0.60 should be the threshold.

- One commenter raised concerns about the potential unintended consequences of endorsing the measure, and the unknown number of truly preventable readmissions.
- Committee Response:
  - . While the measure that was submitted to NQF has an Intra-Class Correlation Coefficient below 0.60, the Committee believes it represents an acceptable benchmark for reliability for measurement of readmissions following a hospitalization for COPD. The Committee concluded that developers' current approach to risk-adjustment and exclusions met the Scientific Acceptability criteria, and were satisfied with the measure's reliability.
  - The Committee recognizes the potential for negative unintended consequences of admissions and readmissions measures and recommends careful monitoring of their implementation. Above all, the Committee agreed that use of these measures should be monitored to ensure they do not inadvertently reduce access to necessary care. The Committee noted the inverse relationship between mortality and readmission for heart failure and recognized the need for careful surveillance and balancing of these measures. The Committee also reiterated its concerns about the need to carefully balance implementation of measures addressing psychiatric readmissions to prevent the risk of higher suicidality.
  - On the other hand, the Committee has noted the desire to understand a patient's need for any subsequent acute care after a hospitalization. In particular, the Committee recognized the need understand if patients are being seen in the Emergency Department after discharge or being placed in observation. The Committee recommends continued work to ensure that the use of readmissions measures does not result in unnecessary or avoidable use of the emergency department or observation status while ensuring that all patients have access to any necessary care. The Committee noted that a number of measures recommended for endorsement in this project could help to balance these concerns, in particular the measures addressing excess days in acute care and population-based admission measures.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2827 PointRight® Pro Long Stay(TM) Hospitalization Measure

Submission | Specifications

Description: The PointRight Pro Long Stay Hospitalization Measure is an MDS-based, risk-adjusted measure of the rate of hospitalization of long-stay patients (aka "residents") of skilled nursing facilities (SNFs) averaged across the year, weighted by the number of stays in each quarter.

Numerator Statement: The numerator for the measure is the sum over four quarters of the counts of hospitalizations of the quarterly denominator populations, where hospitalizations comprise discharges directly from the SNF to an acute care hospital.

The count of hospitalizations excludes discharges from the SNF to LTACHs, IRFs, and psychiatric hospitals, and excludes admissions to acute care hospitals that directly follow a discharge from the SNF to a setting other than an acute care hospital.

However, if a patient is discharged from a SNF directly to an acute care hospital during a quarter at risk, the hospitalization will be counted in the numerator even if the patient was discharged to a setting other than an acute care hospital earlier in that quarter.

Hospitalizations are counted over at-risk intervals of 3 months at a time because this period is long enough to yield nonzero numerators even for SNFs with low rates of hospitalization, yet short enough so that almost all of the denominator population will be present in the facility for all, or almost all, of the period. The latter feature makes

the calculation simpler than if the risk exposure was calculated by days or weeks.Four quarters of denominators and four quarters of numerators are summed to yield the values for the full measure period.

Denominator Statement: The quarterly denominator population consists exactly of those patients present in the SNF on the first day of the quarter (the "snapshot date") who meet the criterion for long stay on that date. The denominator for a quarter is the number of patients in the quarterly denominator population. The denominator for the measure is the sum of the quarterly denominators for the four quarters in the 12 month measure period.

The criterion for a patient's having a long stay is a cumulative length of stay in the facility of more than 100 days as of the snapshot date. The cumulative length of stay of a patient is the length of the current stay as of the snapshot date and plus the full lengths of stay of any previous stays that are linked to it. According to the criteria for linkage of stays used in the present measure, a stay in a SNF is linked to a subsequent stay in the SNF if the patient was discharged from the SNF to the community and was readmitted to the SNF within 10 days or fewer. All stays in a sequence of linked stays are included in the sum of days used to determine a patient's cumulative length of stay. In these criteria the term "community" comprises private residences and all organized settings that are primarily residential in character, including senior housing, independent living facilities, board and care homes, and assisted living facilities.

A patient can contribute multiple times to the denominator for a 12 month measure period. For example, a resident continuously present in the facility for a full year would contribute four to the denominator.

Exclusions: There are no exclusions from the denominator; all patients in the facility on the snapshot date who meet the long stay criterion on that date are included. However, the measure will not be reported for a SNF if the annual unknown outcome rate is greater than 10%.The definition of the annual unknown outcome rate is provided in S.11.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: The risk adjustment model for PointRight Pro Long Stay Hospitalization Rate begins by segmenting the quarterly denominator population for each quarter into four groups based on the duration of the patient's current stay in the SNF. The denominator population is segmented into these four groups because even after controlling for the other risk adjusters, significant variation by length of stay remains and the coefficients within the length of stay groups are different. For each group the risk of one or more discharges from the SNF directly to an acute care hospital during the quarter is estimated by a logistic regression. (Note that the dependent variable is a binary variable rather than the count of hospitalizations of the patient during the quarter.) The independent variables in each logistic regression model come from the patient's most recent MDS 3.0 assessment prior to the snapshot date that has the variable. (Not all of the independent variables in the logistic regressions are present on every type of MDS assessment; this implies that it is sometimes necessary to extract independent variables from two or more discrete MDS assessments.)

The four logistic regression models use subsets of the following set of independent variables. In S.18 below, MDS items corresponding to each listed variable are provided.

Active Diagnoses (A diagnosis is "active" if it affects the patient's current clinical status or treatment plan. An active diagnosis must be documented in the medical record by a physician or physician extender to be checked off in the MDS. Diagnoses are used in the model only if they are indicated in check boxes on Section I of the MDS; if they are indicated by write-in codes in MDS item I8000 they are not utilized in determining the values of the independent variables.):

-Anemia

-Chronic Lung Disease (including Asthma and COPD) -Chronic Lung Disease receiving oxygen therapy at least one time in the 14 days prior to the MDS date

-Diabetes Mellitus receiving insulin at least once in the 7 days prior to the MDS assessment reference date

-Gastroesophageal Reflux Disease (GERD) or Ulcer (esophageal, gastric, or duodenal)

-Heart Failure

-Hypertension

-Viral Hepatitis

2827 PointRight® Pro Long Stay(TM) Hospitalization Measure

-Neurogenic Bladder
-Renal Insufficiency, Renal Failure, or End-Stage Renal Disease

Incontinence:
-Total bowel incontinence

Demographics:
-Age 90 or over
-Male

Medications received at least once within the 7 days prior to the MDS assessment reference date:
-Anticoagulant
-Antibiotic

Context of Care:
-Current stay began with admission from an acute care hospital
-In this SNF 6 months before the snapshot date (whether or not in the facility continuously for the 6 months preceding the snapshot date
-In this SNF 12 months before the snapshot date (whether or not in the facility continuously for the 12 months preceding the snapshot date
-Natural log of (the length of the current stay as of the snapshot date minus 100 days). (Linked stays are not included in this calculation.)

Symptoms:
-Dyspnea (shortness of breath or trouble breathing) on exertion

Skin condition:
-Surgical wound(s)

Hospice Status:
-Receiving hospice care while resident in the facility, at some time during the 14 days prior to the MDS assessment reference date

Treatments (given in the facility at least once in the 14 days preceding the MDS assessment reference date):
-IV fluid or medication
-Oxygen therapy

Socioeconomic Status:
- Medicaid beneficiary (as indicated by having a Medicaid number or having a Medicaid number pending)
- Black or African-American race/ethnicity (as described the patient or family, either as a sole identity or one of several, e.g., black and Caucasian, black and Latino)

Level of Analysis: Facility
Setting of Care: Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility
Type of Measure: Outcome

2827 PointRight® Pro Long Stay(TM) Hospitalization Measure

Data Source: Electronic Clinical Data

Measure Steward: American Health Care Association

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-18; N-0**; 1b. Performance Gap: **H-7**; **M-11**; **L-0**; **I-0**

<u>Rationale</u>:

- As a rationale for measuring this health outcome, the developer suggests that skilled nursing facilities are able to influence rates of hospitalizations for long term care residents in an number of ways including structural interventions such as high staffing levels and nurse practitioner availability as well as process interventions such as early detection of signs and symptoms of impending infections (pneumonia, urinary tract infection, etc.) and chronic disease exacerbation (e.g. congestive heart failure, diabetes mellitus, etc.)
- The developer cited a 2010 study showing that 33% of SNFs hospitalization can be avoidable, and in 2005 (according to the same 2010 study), avoidable hospitalizations cost Medicare $3 billion and Medicaid $463 million. Additionally, the developer presented data obtained from the national MDS data from CMS, citing 437,356 long nursing home stays discharged to an acute hospital from the first quarter of 2015.
- The Standing Committee discussed the need for this measure, noting the lack of measures for this population, as well as the need to identify and study hospitalizations among long stay residents. The Standing Committee noted the current focus on short-term stay patients, rather than long-stay. The fact that many hospitalizations of this population can often be avoided (between 25% to 33% as stated by the Standing Committee), further emphasized the importance of this measure.
- The Standing Committee agreed the measure met the evidence criteria.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-4**; **M-14**; **L-0**; **I-0** 2b. Validity: **H-1**; **M-18**; **L-0**; **I-0**

<u>Rationale</u>:

- The developers performed three types of reliability testing including alignment of model independent values, reliability of rates over time, and the stability of facility level adjusted rate bootstrapping.
    - The developers compared the prevalence of the risk adjustment covariates between a testing sample of 2,096 SNFs and the national population and analyzed change from quarter to quarter in the observed and adjusted long-stay hospitalization rates.
- The developer explained that their reasoning was that the underlying probability of a SNF's long-stay patients hospitalizing and the characteristics of its long-stay patient population were unlikely to change greatly in a three month period so that most of the change from quarter to quarter would be due to limitations on measure reliability.
- The developer recalculated adjusted rates for the measure for CY 2014 using a random sample of stays. The developer then reviewed the distribution of differences between facilities' original adjusted rates and the rates calculated with the new sample. The developer interpreted a distribution of differences with a small variance and a mean of zero as acceptable measure stability or reliability.
    - The developer interpreted their results as representative of the SNF population and 48% of the comparable risk adjustment model covariates were found to have prevalence within 5% of the prevalence found in the national sample. 66% were found to have prevalence within 10% of the prevalence found in the national sample.
- The developer performed two methods of validity testing including agreement of model dependent variables, and the performance measure score in correlation with the SNF industry measures of quality.
    - The comparison showed that that 86% of hospitalizations of Medicare FFS patients identified by the MDS are confirmed by Medicare FFS claims; in the other direction, 98% (208,891 of 213,772) of acute inpatient claims found near an MDS discharge have an MDS discharge code of acute hospital.

2827 PointRight® Pro Long Stay(TM) Hospitalization Measure

  - The developer interprets this finding that MDS discharge assessments appear to be overstating the rate of acute hospitalizations to a moderate degree but that the overall high level of agreement between MDS discharge coding and claims supports the validity of the measure
- The differences between age and race categories were noted by the Standing Committee during the validity discussion. Although the developers noted discharge to community rates as well as other negative outcomes that differ by race and age, the Standing Committee noted that the measure itself is separate from these issues. The Standing Committee agreed the developer had provided a conceptual reason not to include the small effects identified.
  - The developer stated that race was included because of the Standing Committee discussion from the year prior. However, upon further inspection and discussion with the developer, the Standing Committee requested that race be removed from the measure. Under instruction from NQF staff, the Standing Committee continued voting on the measure under the assumption that the developer would remove race at later date.
  - The developers have since updated the measure to remove race from the risk adjustment model.
- The risk adjustment model employed in the PointRight Pro Long Stay Hospitalization Rate utilizes four logistic regression models applied to four discrete subgroups of the denominator population to estimate risk of any hospitalization during a quarter at risk.
  - Logistic Regression Model Long Stay Group 1 c-statistic = .63
  - Logistic Regression Model Long Stay Group 2, c-statistic = .63
  - Logistic Regression Model Long Stay Group 3, c-statistic = .62
  - Logistic Regression Model Long Stay Group 4, c-statistic = .63
  - Linear Regression Model Rate of all Hospitalizations, R-squared = .99
- The Standing Committee also had questions about the dataset, but the developer confirmed that the measure is based on the Minimum Data Set (MDS) and therefore it is not based on claims data.
- The Standing Committee agreed this measure met the reliability and validity criteria.

3. Feasibility: H-12; M-7; L-0; 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 required data elements are routinely generated and used during care delivery. They are collected and used by healthcare personnel during the provision of care.
- Although some Standing Committee Members noted the burden that this measures can cause for a nursing home staff because of the changes that would likely results from the use of this measure, such as changing staffing patterns, the Standing Committee agreed the measure would be feasible and worth the effort.

4. Usability and Use: H-4; M-14; L-1; I-0

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

Rationale:

- This measure is not currently reported but is planned for use in CMS' evaluation of SNF's clinical performance. Also, AHCA plans to publish this measure on its website for free use by AHCA members and other selected stakeholders. The Standing Committee raised the issue discussed under feasibility and the fact that effort would be required by nursing homes under this measure, but the Standing Committee agreed the measure was usable.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-18**; **N-1**

Rationale

- The Standing Committee recognized the importance of reducing the number of home health stays for patients who have a Medicare claim for an unplanned admission to an acute care hospital in the 60 days

2827 PointRight® Pro Long Stay(TM) Hospitalization Measure
following the start of the home health stay and recommended the measure for endorsement.
6. Public and Member Comment • Two comments were received on this measure. Commenters agreed with the endorsement of the measure but raised concerns about its potential application at the health plan level as it uses electronic clinical data that is not feasible for plans to collect. • Committee response: ○ The Committee agreed that the measure should be applied at the facility-level, as it is specified and tested. The Committee believes that linking claims and EHR data is an important advancement in quality measurement.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2858 Discharge to Community
Submission | Specifications
Description: The Discharge to Community measure determines the percentage of all new admissions from a hospital who are discharged back to the community alive and remain out of any skilled nursing center for the next 30 days. The measure, referring to a rolling year of MDS entries, is calculated each quarter. The measure includes all new admissions to a SNF regardless of payor source. **Numerator Statement**: The outcome measured is the number of new admissions from an acute care hospital discharge to community from a skilled nursing center. More specifically, the numerator is the number of stays discharged back to the community (i.e. private home, apartment, board/care, assisted living, or group home as indicated on the MDS discharge assessment form) from a skilled nursing center within 100 days of admission and remain out of any skilled nursing center for at least 30 days. **Denominator Statement**: The denominator is the total number of all admissions from an acute hospital (MDS item A1800 "entered from"=03 (indicating an "acute care hospital") to a center over the previous 12 months, who did not have a prior stay in a nursing center for the prior 100 days (calculated by subtracting 100 from the admission date (MDS item A1900 "admission date"). Please note, the denominator only includes admissions from acute hospitals (MDS item A1800 "entered from"=03 (indicating an "acute care hospital") regardless of payor status. **Exclusions**: The denominator has three exclusions (see below). First, stays for patients less than 55 years of age are excluded from the measure. Second, stays for which we do not where the patient entered from, or for which we do not observe the patient's discharge, are excluded from being counted in the denominator. Third, stays with no available risk adjustment data (clinical and demographic characteristics listed in Section S.14) on any MDS assessment within 18 days of SNF admission are excluded from the measure. Note, while not denominator exclusions, we also suppress the data for facilities that have fewer than 30 stays in the denominator, or for whom the percent of stays with a known outcome is less than 90%. The suppression of risk adjusted to community rates for facilities with fewer than 30 stays in the denominator is to improve the reliability of the measure, as detailed in the testing section (2b3). The suppression of rates for facilities for whom fewer than 90% of stays had a known outcome is done to improve the reliability of the measure and avoid perverse incentives about submitting MDS assessments for patients not discharged to the community. **Adjustment/Stratification**: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Risk adjustment for the measure was completed by means of logistic regression using independent variables drawn from the admission to SNF and discharge from SNF MDS 3.0 assessments. When information was not available on the admission MDS assessment, information from the next available MDS of any type (except discharge MDS assessment) was used, as long as the MDS was completed within 18 days of admission to the center; if no such complete assessment exists on entry or within 18 days, the stay is excluded from the denominator per the

2858 Discharge to Community

denominator exclusions.

The following lists the variables used in the logistic regression risk adjustment model. There are 60 different MDS items, which are encoded across 116 variables in the final risk model (e.g., age and age-squared; interaction terms; etc.). The respective MDS 3.0 codes used to determine whether or not each variable contributes to the calculation are provided in Section S.15 below.

Demographic:
-Age
-Gender
-Marital Status

Functional Status:
-Vision
-Makes Self-understood
-Ability to Understand

Functional Status (cognitive, mobility and self care):
-Any Sign/symptom of Delirium
-Major Depression
-Behavioral Code (i.e. Hallucination, Delusion, Physical Behavior, Verbal Behavior, Other Behavior)
-Any Rejection of Care
-Medicare RUG IV Hierarchical Group
-Activities (i.e Bed Mobility, Transfer, Walk in Corridor, Locomotion, Eating, and Personal Hygiene)
-ADL summary (Combination of Bed Mobility, Transfer, Locomotion, Dressing, Eating, Toilet Use, Hygiene)
-ADL*Cognitive Impairment: Interaction Term
-Bathing
-Balance (i.e. Moving from Seated to Standing, Walking, Turning Around and Facing the Opposite Direction, and Moving On and Off Toilet)
-Urinary Incontinence
-Bowel Incontinence

Prognosis:
-Any acute Hospitalization within 30 days of Admission
-Special Treatment/Programs: Hospice Post-Admission
- Life Expectancy of less than 6 months

Clinical Conditions:
-Shortness of Breath when Exertion
-Shortness of Breath when Sitting
Shortness of Breath when Lying Flat
-Any Swallowing Disorder
-Weight Loss
-Pressure Ulcer
-Wound Infection

2858 Discharge to Community

-Hemiplegia
-Paraplegia

Clinical Treatments:
-Oxygen Post-admit
-Tracheostomy Post-admit
-Ventilator Post-admit
-Dialysis Post-admit
-Max Number Injections
-Antipsychotic Use

Clinical Diagnosis:
-Anemia
-Heart Failure
-Hypertension
-Pneumonia
-Septicemia
-Urinary Tract Infection (UTI)
-Viral Hepatitis
-Diabetes Mellitus
-Hyperkalemia
-Hyperlipidemia
-Hip Fracture
-Other Fracture
-Alzheimer's Disease
-Stroke
-Dementia
-Huntington's
-Malnutrition
-Anxiety Disorder
-Depression
-Manic Depression
-Psychotic
-Schizophrenia
-Asthma, COPD, Chronic Lung Disease

Level of Analysis: Facility
Setting of Care: Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility
Type of Measure: Outcome
Data Source: Electronic Clinical Data
Measure Steward: American Health Care Association

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]
1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>
(1a. Evidence, 1b. Performance Gap)
1a. Evidence: **Y-19; N-0**; 1b. Performance Gap: **H-7**; **M-12**; **L-0**; **I-0**

2858 Discharge to Community

Rationale:

- The developer stated the rationale for the measure that improving national discharge to community rates directly aligns with NQS 3 aims of Better Care, Healthy People/Health Communities, and Affordable Care. The developer listed several studies from peer-reviewed journals that provide examples of clinical actions (identifying warning symptoms, medication reconciliation, follow-ups on labs and appointments, etc.) especially continuous communication between the patient/his family, staff at acute care hospitals and SNF staff lead to a patient- and family-centered improvement of quality of care.
- Studies show the majority of nursing home residents prefer community discharge over remaining in post-acute and long-term care but an estimated 10%-20% of nursing home residents capable of successfully residing in the community with appropriate rehabilitative services and support in place do not get discharged and remain unnecessarily in institutionalized care.
- Extended SNF stays increase a patient's risk and exposure to health care-related infections and serious illnesses, such as Methicillin-resistant Staphylococcus aureus (MRSA) and clostridium difficile (C. difficile). Approximately 2 million infections occur in nursing homes each year (Strausbaugh & Joseph, 2000). Nearly 10-30% of nursing home residents are colonized with C. difficile at any given time (Makris & Gelone, 2007).
- The utilization of SNFs and discharge to community rates is not uniform across the nation or between communities. Non-uniform rates are reflective of inconsistent community practices and engagement in the SNF discharge to community process.
- The Standing Committee specifically noted the importance of this measure since it is the most direct signals of the policy objective to address discharge coordination planning. The Standing Committee noted the relationship this measure has to the ACHA Quality Initiative goal and the importance of measuring the discharge to community rates for skilled nursing facility patients.
- The Standing Committee noted that ten to 20 percent of nursing home residents that are capable of going back to the community remain institutionalized and reference exposure to health care associated infections, as well as psychosocial and financial challenges these residents may experience. The Standing Committee agreed this measure met the evidence criteria.

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-3**; **M-16**; **L-0**; **I-X** 2b. Validity: **H-0**; **M-19**; **L-0**; **I-0**

Rationale:

- The developers used a replacement bootstrapping method and performance comparison between quarters to test for reliability.
- The developer conducted a random resampling of the population with replacement to simulate a facility or two facilities of similar size independently drawing patients from the same underlying patient population and compared outcomes before and after resampling.
    - It was found that if a SNF's patients were completely redrawn from the same underlying population (e.g. the same SNF a year in the future) or if two SNFs who each drew patients from the same underlying population were compared, 68% of the time they will remain ranked within ten percentiles of where they were before redrawing patients. In 96% of cases, they would shift less than thirty percentiles after random resampling.
- The developer tested the validity of the measure two ways. First, the coding of discharges was validated against matched Part A claims data. Secondly, the developer performed construct validity testing by correlating risk adjusted discharge to community rates with certain other measures hypothesized to be driven by the same factors driving discharge to community rates.
    - The developers found a negative and statistically significant relationship between the discharge to community rate and the short stay rehospitalization rate (Pearson's correlation =-0.092, p<.0001).
    - The developer noted this negative correlation was expected because higher scores of discharge to community measure are indicative of higher quality, whereas lower scores of the short stay rehospitalization rate are indicative higher quality.

2858 Discharge to Community

- The developer also found statistically significant correlations between the discharge to community rate and the CMS Nursing Home Compare Short Stay quality measures. These findings were interpreted as supporting the construct validity of the discharge to community measure.
- The risk adjustment model includes 60 risk adjustment variables, which were encoded in 116 variables in the final risk model (including interaction terms, multilevel factor variables, etc.).
- SDS variables were analyzed in the same way as all other variables. The developer did not do any separate analyses on these variables.
- Ultimately the developers included age, sex, and marital status.
- The C-statistic was 0.820.
- The Standing Committee noted the two separate reliability methods used, include replacement bootstrapping and performance comparison. The Standing Committee agreed the measure met the reliability criterion.
- The Standing Committee noted that the measure was adjusted for age, gender, and marital status. The developers noted the correlation between discharges to the community and higher quality care, which was generally agreed upon by the Standing Committee. The Standing Committee agreed the measure met the scientific acceptability criteria.

3. Feasibility: H-16; M-3; L-0; 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:

- All measure data elements are in defined fields in electronic claims (e.g., clinical registry, nursing home MDS, home health OASIS) and routinely collected by and used by healthcare personnel during the provision of care. It was determined that this measure did not present collection burden because it relies solely on data items from the MDS 3.0 that all facilities are already required to submit.
- The Standing Committee noted that there would likely be fluctuation between quarter to quarter due to missing rates, but overall agreed the measure was feasible.

4. Usability and Use: H-4; M-15; L-0; I-0

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

Rationale:

- This measure is currently publicly reported and is used in an accountability program. The measure is currently used for quality improvement and benchmarking.
- The measure has been in use since 2014 and the Standing Committee noted a 3.6% increase. The Standing Committee made a suggestion to provide more clarification in the title, but overall agreed the measure would be highly usable.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-19**; **N-0**

Rationale

- The Standing Committee recognized the importance of improving national discharge to community rates and recommended the measure for endorsement.

6. Public and Member Comment

- Two comments were received on this measure. Commenters agreed with the endorsement of the measure but raised concerns about its potential application at the health plan level as it uses electronic clinical data that is not feasible for plans to collect.
- Committee Response:
  - The Committee agreed that the measure should be applied at the facility-level, as it is specified and tested. The Committee believes that linking claims and EHR data is an important

2858 Discharge to Community
advancement in quality measurement.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF)

Submission | Specifications

Description: This facility-level measure estimates an all-cause, unplanned, 30-day, risk-standardized readmission rate for adult Medicare fee-for-service (FFS) patients with a principal discharge diagnosis of a psychiatric disorder or dementia/Alzheimer's disease.

The performance period for the measure is 24 months.

Numerator Statement: The measure estimates the incidence of unplanned, all-cause readmissions to IPFs or short-stay acute care hospitals following discharge from an eligible IPF index admission. We defined readmission as any admission that occurs on or between Days 3 and 30 post-discharge, except those considered planned.

Denominator Statement: The target population for this measure is Medicare FFS beneficiaries aged 18 years and older discharged from an inpatient psychiatric facility with a principal diagnosis of a psychiatric disorder. Eligible index admissions require enrollment in Medicare Parts A and B for 12 months prior to the index admission, the month of admission, and at least 30 days post discharge. Patients must be discharged alive to a non-acute setting (not transferred). A readmission within 30 days is eligible as an index admission, if it meets all other eligibility criteria.

Exclusions: The measure excludes admissions for patients:

- Discharged against medical advice (AMA)
- With unreliable data (e.g. has a death date but also admissions afterwards)
- With a subsequent admission on day of discharge and following 2 days (transfers/interrupted stay period)

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Hierarchical logistic regression is used to estimate a risk standardized readmission rate.

CANDIDATE AND FINAL RISK FACTOR VARIABLES

Four types of risk factors were considered based on empirical analysis, literature review, and clinical judgment:

1. Principal discharge diagnosis of the IPF index admission: Discharge diagnoses were summarized into 13 distinct principal discharge risk variables using a modified version of AHRQ CCS.
2. Comorbidity risk variables: Identified from secondary diagnoses of the index admission and primary or secondary diagnoses of in- and outpatient encounters during the 12-month look-back period using modified CMS condition categories (CC)
3. Other risk factors variables from literature such as history of discharge AMA, aggression and self-harm
4. Age and gender

FINAL SET OF RISK-ADJUSTMENT VARIABLES

Age (7 levels), gender

Principal discharge diagnoses (13)

CCS 650 Adjustment disorder

CCS 651 Anxiety

CCS 652/654/655 ADD/Developmental/Childhood disorders

CCS 653 Dementia

2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF)

CCS 656 Impulse control disorders
CCS 657.1 Bipolar disorder
CCS 657.2rc Depressive disorder
CCS 658 Personality disorder
CCS 659.1 Schizo-affective disorder
CCS 659.2 Psychosis
CCS 660 Alcohol disorder
CCS 661 Drug Disorder
CCS 670/663 Other mental disorder
Comorbidities: 26 non-psychiatric CC, 12 psychiatric CC groups
CC Description (CC or ICD-9-CM)
AMI (CC 81, 82)
Anemia (CC 47)
Arrhythmia (CC 92, 93)
Asthma (CC 110)
COPD/Fibrosis (CC 108, 109)
Delirium (CC 48)
Diabetes (CC 19, 119, 120)
Diabetes complications (CC 15-18)
Dialysis (CC 130)
Endocrine disease (CC 22, 23)
Heart disease (CC 83, 84, 89, 90, 104-106)
Heart failure (CC 80)
Hematological disorder (CC 44)
Infection (CC 1, 3-5, 37, 152)
Injury (CC 150, 151, 155, 156, 160, 162, 163)
Liver disease (CC 25-29)
Lung problems (CC 111-115)
Malnutrition (CC 21)
Metastasis (CC 7)
Organ transplant (CC 174, 175)
Other infection (CC 6)
Pancreatic disease (CC 32)
Peptic ulcer (CC 34)
Seizures (CC 74)
Uncompleted pregnancy (CC 142, 146, 147)
Urinary tract disorder (CC 136)
Adjustment disorder (ICD-9-CM 309.0, 309.22-309.24, 309.28-309.29, 309.3-309.4, 309.82-309.83, 309.89, 309.9, 309.1)
Anxiety (ICD-9-CM 293.84, 300.01-300.02, 300.00, 300.09, 300.10, 300.20-300.23, 300.29, 300.3, 300.5, 313.0, 313.21, 313.22)
Bipolar (ICD-9-CM 296.00-296.06, 296.10-296.16, 296.40-296.46, 296.50-296.56, 296.60-296.66, 296.7, 296.80-296.82, 296.89, 296.90, 296.99)

2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF)

Depression (ICD-9-CM 296.20-296.26, 296.30-296.36, E950.0-951.1, E951.8, E952.0-952.1, E952.8-953.1, E953.8-953.9, E954, E955.0-955.7, E955.9, E956, E957.0-957.2, E957.9-958.9, E959, 300.4, 311, V62.84)

Developmental disability (CC 66 + ICD-9-CM 758.6-758.7, 758.81, 758.89, 758.9, 759.4, 759.89, 313.1, 313.3, 313.81-313.83, 315.00-315.02, 315.09, 315.1-315.2, 315.31-315.32, 315.34-315.35, 315.39, 315.4-315.5, 315.8-315.9, 313.23, 313.89, 313.9)

Drug/alcohol disorder (CC 51, 52, 53 (except ICD9-CM 305.1) + ICD-9-CM CM 648.31-648.32, 648.34, 655.51, 648.30, 648.33, 655.50, 655.53, 980.0, 965.00-965.02, 965.09, 760.71-760.73, 760.75, 779.5, v654.2)

Intellectual disability (CC 61-64)

Other psych disorders (ICD-9-CM 300.11-300.13, 300.15-300.16, 300.19, 300.6-300.7, 300.81-300.82, 307.1, 307.51, 799.2, 799.21-799.25, 799.29, 300.89, 300.9, 308.0-308.4, 308.9, 312.8, 312.00-312.03, 312.10-312.13, 312.20-312.23, 312.4, 312.81-312.82, 312.89, 312.9, 307.0, 307.9, 307.20-307.23, 307.3, 307.6, 307.7, 309.21, 312.30-312.35, 312.39, 302.0-302.4, 302.50-302.53, 302.6, 302.70-302.76, 302.79, 302.81-302.85, 302.89, 302.9, 306.0-306.4, 306.50-306.53, 306.59, 306.6-306.9, 307.40-307.50, 307.52-307.54, 307.59, 307.80, 307.89, 316)

Personality disorder (CC 57)

Psychosis (CC 56 + ICD-9-CM 295.00-295.05, 295.10-295.15, 295.20-295.25, 295.30-295.35, 295.40-295.45, 295.50-295.55, 295.60-295.65, 295.80-295.85, 295.90-295.95, 297.0-297.3, 297.8-297.9)

PTSD (ICD-9-CM 309.81)

Schizo-affective (ICD-9-CM 295.70-295.75)

Discharged AMA in prior 12 months

Suicide attempt/self-harm — identified by the presence of at least one inpatient or outpatient claim with diagnosis of suicidal attempt or self-harm in the 12-month look-back period.

Aggression — identified by the presence an ICD-9-CM code indicating aggression as a secondary diagnosis on the index admission or on an inpatient or outpatient claim in the 12-month look-back period.

Level of Analysis: Facility

Setting of Care: Behavioral Health/Psychiatric : Inpatient

Type of Measure: Outcome

Data Source: Administrative claims

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-18; N-1**; 1b. Performance Gap: **H-8**; **M-10**; **L-1**; **I-0**

<u>Rationale</u>:

- An analysis of Medicare claims data found that over 20% of patients who receive psychiatric care in an inpatient setting are readmitted within 30 days of discharge.
- The Standing Committee stressed the evidence that readmission rates can be lowered though care coordination interventions and discharge planning practices, such as improving care management and connecting patients to services in their communities. The Standing Committee noted that a lack of care coordination is an on-going issue in behavioral health and that rates of connection with aftercare following discharge from an inpatient facility are low.
- The measure developer provided the distribution of 11.0% to 35.4% with an average rate of 21.0%
- The Standing Committee noted there is a need for an increased focus on admissions, readmissions, and care coordination issues in behavioral health. In particular, the Standing Committee noted that the limited data available suggests readmissions for behavioral health may be higher than general medical/surgical readmissions and that there are currently very low rates of connections to aftercare.
- The Standing Committee agreed that there were interventions such as intensive care management and

2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF)

connections to services in the community that could improve the results of this measure.
- The Standing Committee noted unique challenges in the behavioral health setting and raised concerns about the impact of access to care on this measure. The Standing Committee raised concerns that this measure should be implemented carefully to avoid worsening access issues.
- Based on these results the Standing Committee concurred a gap in care exists and that there is an opportunity for improvement.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-3**; **M-16**; **L-0**; **I-0** 2b. Validity: **H-0**; **M-16**; **L-4**; **I-0**

<u>Rationale</u>:

- To test the reliability of the measure, the developer calculated the intra-class correlation coefficient (ICC) using a test-retest approach that examines the agreement between repeated measures of the same IPF for the same time period.
- The developer used two test-retest approaches to generate independent samples of patients within the same IPF: a split-half sampling design and bootstrapping.
    - For split-half sampling, the developer randomly sampled half of all eligible index admissions in each facility over the two-year period, resulting in two samples that cover the same two-year period but with case volume the size of a measure that would be calculated with one year of data. The ICC in the split-half sampling design was estimated using the RSRRs of the two split-half samples.
    - A total of 716,174 admissions over a 2-year period were examined, with 358,087 in each randomly-selected sample. The RSRR was estimated for each sample using a hierarchical logistic regression model. The average RSRR in the two-split-half samples had means of 21.03% and 20.93 percent. The agreement between the two RSRRs for (as measure by an intra-class correlation coefficient (ICC)) was 0.60.
    - For bootstrapping, the developer sampled 1,000 pairs of samples from the original measure cohort with replacement (stratified sampling by IPF), resulting in 1,000 pairs of new samples within each IPF with the identical sample size as in the original measure cohort, thus maintaining the sample size of a two-year measure. The ICC in the bootstrap sampling was estimated for each pair of the bootstrap samples. With the 1,000 ICC estimates from the 1,000 pairs of bootstrap samples, the developer determined the distribution of estimated ICC coefficients and thus could calculate the mean and 95% CI of the ICC.
    - The ICC obtained from the bootstrapping approach, comparing 1,000 pairs of samples of the original measurement cohort, which were sampled with replacement yielding an identical sample size as the original measurement cohort, is 0.78 (95% CI 0.77-0.80).
- The developer performed a systematic assessment of face validity of the measure score. Face validity of the measure score was obtained by a TEP vote at the conclusion of measure development.
- This measure employs a hierarchical logistic regression model (a form of hierarchical generalized linear model [HGLM]) to create a hospital level 30-day risk-standardized readmission rate (RSRR). To validate the risk adjustment model, the developer used bootstrapping in which 1,000 bootstrap samples were randomly drawn from the original dataset with replacement. The bootstrap samples were used as the development dataset, and the original cohort was used as the comparison dataset. The C-statistic was 0.660.
- To select clinical risk factors, the developers employed a stepwise logistic regression process with backward elimination of variables, using 100 bootstrap samples derived from the entire measure population via random selection with replacement. The developer retained all variables in the stepwise backward elimination that showed an association with readmission at p<.15 in 70% of the bootstrap samples.
- The developers also considered a number of variables related to sociodemographic status (SDS) for

2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF)

potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources. Given the complexity of accurately measuring SDS in current datasets, the developers do not think the empirical evidence is strong enough to warrant inclusion of any of the current SDS variables in the risk model for this measure.

- The Standing Committee raised concerns about the number of patients being excluded because of transfers and interrupted stays. In particular, the Standing Committee raised concerns that this excludes the sizeable number of patients who are discharges from an IPF but readmitted a day or two later. However, the Standing Committee recognized that at this time it is not possible to capture this data from Medicare claims. The Standing Committee expressed a desire to see this issue explored further in the future.
- The Standing Committee raised concerns about the 24 month timeframe for this measure but accepted the developer's rationale that this would allow more facilities to achieve the minimum threshold of 25 cases.
- The Standing Committee urged the developer to consider ways to expand the measure beyond Medicare patients.
- The Standing Committee agreed this measure met the reliability and validity criteria.

3. Feasibility: H-12; M-8; L-0; 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:

- This measure is collected through administrative claims data.
- The Standing Committee agreed the measure would be feasible to collect and implement.

4. Usability and Use: H-6; M-13; L-1; I-0

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

Rationale:

- The measure is not currently publicly reported, but it is intended for use in the Inpatient Psychiatric Facility Quality Reporting Program.
- The Standing Committee noted the need to be able to measure readmissions in behavioral health, however the Standing Committee recognized the challenges of patient engagement in this population
- The Standing Committee did express concerns about the unintended consequences of this measure, in particular they noted the need to protect access to care and to balance this measure with measures addressing outcomes like mortality.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-19**; **N-1**

Rationale

- The Standing Committee recognized the importance of reducing readmissions to inpatient psychiatric facilities and the need for improved care coordination and discharge management. The Standing Committee noted the unique challenges of connecting with follow-up care in behavioral health and also noted the need to monitor other outcomes in this population such as mortality.

6. Public and Member Comment

- One comment was received in support of endorsement of this measure.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-X; N-X

2860 Thirty-day all-cause unplanned readmission following psychiatric hospitalization in an inpatient psychiatric facility (IPF)

9. Appeals

2879 Hybrid Hospital-Wide Readmission Measure with Claims and Electronic Health Record Data

Submission | Specifications

Description: The measure estimates a hospital-level risk-standardized readmission rate (RSRR) of unplanned, all-cause readmission after admission for any eligible condition within 30 days of hospital discharge. The measure reports a single summary RSRR, 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. The measure also indicates the hospital-level standardized readmission ratios (SRR) for each of these five specialty cohorts. The outcome is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission (the admission included in the measure cohort). A specified set of planned readmissions do not count in the readmission outcome. The target population is Medicare Fee-for-Service beneficiaries who are 65 years or older.

This Hybrid Hospital-Wide Readmission (HWR) measure is a re-engineered version of measure 1789, the Hospital-Wide All-Cause Unplanned Readmission Measure which was developed for patients 65 years and older using Medicare claims and is currently publically reported in the Hospital Inpatient Quality Reporting Program. This reengineered measure uses clinical data elements from patients' electronic health records in addition to claims data for risk adjustment.

Numerator Statement: The outcome for this measure is 30-day readmission. We define readmission as an inpatient admission for any cause, with the exception of certain planned readmissions, within 30 days from the date of discharge from an eligible index admission. If a patient has more than one unplanned admission (for any reason) within 30 days after discharge from the index admission, only one is counted as a readmission. The measure looks for a dichotomous yes or no outcome of whether each admitted patient has an unplanned readmission within 30 days. However, if the first readmission after discharge is considered planned, any subsequent unplanned readmission is not counted as an outcome for that index admission because the unplanned readmission could be related to care provided during the intervening planned readmission rather than during the index admission.

Denominator Statement: The measure includes admissions for Medicare beneficiaries who are 65 years and older and are discharged from all non-federal, acute care inpatient US hospitals (including territories)with a complete claims history for the 12 months prior to admission.

Additional details are provided in S.9 Denominator Details.

Exclusions: The measure excludes index admissions for patients:

1. Admitted to Prospective Payment System (PPS)-exempt cancer hospitals;
2. Without at least 30 days post-discharge enrollment in FFS Medicare;
3. Discharged against medical advice (AMA);
4. Admitted for primary psychiatric diagnoses;
5. Admitted for rehabilitation; or
6. Admitted for medical treatment of cancer.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

2879 Hybrid Hospital-Wide Readmission Measure with Claims and Electronic Health Record Data

Data Source: Administrative claims, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Laboratory

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

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

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

1a. Evidence: **Y-19; N-0**; 1b. Performance Gap: **H-2**; **M-17**; **L-0**; **I-0**

Rationale:

- This hybrid measure estimates a hospital-level risk-standardized readmission rate (RSRR) for unplanned readmission for any eligible condition within 30 days of hospital discharge, using both claims and electronic health record data (EHR). Electronic clinical information is added into the risk adjustment model to enhance the face validity and performance of the measure.
- The Standing Committee noted that the while the opportunity for improvement on this measure may be the same as #1789, the inclusion of clinical data through hybrid measures is an opportunity for innovation for future measures and could improve and enhance quality measurement.

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-17**; **L-1**; **I-0** 2b. Validity: **M-16**; **L-2**; **I-0**

Rationale:

- Reliability testing was performed at both the measure score and data element levels. The measure was developed to avoid the use of claims data elements through to be coded inconsistently across hospitals, instead using filed that are consequential for payment and which are audited by CMS. In addition, the developer compared frequencies and odds ratios of variables from their risk model across three years of data in order to assess the consistency of those variables over time. The performance score was assessed through test-retest reliability. The agreement between the two RSRRs for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.688.
- The Standing Committee noted that the developers used Health Quality Measure Format (HQMF) specifications and used the Value Set Authority Center (VSAC) for their code sets. Additionally, the measure was created using the measure authoring tool (MAT). The use of these tools should help to ensure this measure can be implemented reliability.
- However, the Standing Committee expressed concerns about the reliability of EHR data and that the measurement error associated with EHRs is going to be different from measurement error associated with claims data.
- The validity of the measure was assessed through face validity. The measure was tested at both the measure score and data element levels.
- The measure employs a hierarchical logistic regression model (a form of hierarchical generalized linear model [HGLM]) to create a hospital-level 30-day risk-standardized readmission rate (RSRR).
- Several critical clinical data elements used in the measure's risk models were derived from patients' electronic medical records. When this measure is implemented, CMS intends to obtain these critical data elements from hospital EHRs and merge the data with claims data to calculate and report measure results.
- The developer tested the validity of electronic extraction of these critical data elements as part of a more comprehensive evaluation of a larger set of core clinical data elements (CCDEs). The CCDE are a set of 21 EHR data elements that are captured on most adults (plus Troponin, which is a condition-specific CCDE for patients with acute myocardial infarction) admitted to acute care hospitals, are easily extracted from EHRs, and can be used to risk adjust hospital outcome measures for a variety of conditions and procedures. All of the critical data elements used in the Hybrid HWR measure are included in the CCDE.
- The addition of electronic clinical data results in a small improvement in risk model discrimination.
- The developer tested the impact of SDS variables on the risk model. The developer ultimately chose not to include these variables in the model because the effect size of each of these variables is small, the c-

2879 Hybrid Hospital-Wide Readmission Measure with Claims and Electronic Health Record Data

statistic (i.e., predictive value) is unchanged with the addition of any of these variables into the model, and the addition of any of these variables into the model has little to no effect on hospital performance.

- C-statistic for each co-hort:
  - Medicine cohort: 0.651
  - Surgery/Gynecology cohort: 0.802
  - Cardiorespiratory cohort: 0.668
  - Cardiovascular cohort: 0.731
  - Neurology cohort: 0.708

3. Feasibility: H-2; M-16; L-0; 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:

- This measure is based on administrative claims data and electronic clinical data, which will be collected from hospitals using MAT output and value sets to inform data queries and electronic reporting requirements.

4. Usability and Use: H-6; M-11; L-1; I-0

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

Rationale:

- This measure is intended for implementation in the Hospital Inpatient Quality Reporting (IQR) program. The Standing Committee noted that if the data needed to calculate this measure can be feasibly reported it is useful for that purpose.

5. Related and Competing Measures

- This measure directly competes with NQF #1789 Hospital-Wide All-Cause Unplanned Readmission Measure (HWR). The measure estimates a hospital-level risk-standardized readmission rate (RSRR) of unplanned, all-cause readmission after admission for any eligible condition within 30 days of hospital discharge. The measure reports a single summary risk-standardized readmission rate (RSRR), 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. The measure also indicates the hospital-level standardized risk ratios (SRR) for each of these five specialty cohorts. The outcome is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission. Admissions for planned procedures that are not accompanied by an acute diagnosis do not count as readmissions in the measure outcome. The target population is patients 18 and over. CMS annually reports the measure for patients who are 65 years or older and are enrolled in fee-for-service (FFS) Medicare and hospitalized in non-federal hospitals.
- The Standing Committee justified including both measures in the portfolio because #2879 includes additional clinical variables in the risk adjustment model and these additional variables are obtained through EHR data. Due to the current challenges of collecting and reporting EHR data the Standing Committee felt that #2879 may not be ready for wide scale implementation and that both measures should be endorsed.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-2**

Rationale

- The Standing Committee noted that this measure represented an important improvement to quality measurement. Linking claims and electronic clinical data could allow for the inclusion of important new variables in risk adjustment models. However, the Standing Committee recognized the challenges to using and reporting EHR data and to using a measure across EHR systems. The Standing Committee felt that this

hybrid measure offered increased risk model discrimination over the claims-based version (NQF #1789) making it suitable for endorsement.

6. Public and Member Comment

- This measure received four comments. Two comments raised concerns about the data limitations that currently exist for electronic health records.
- One commenter raised concerns about the potential unintended consequences of endorsing the measure, and the unknown number of truly preventable readmissions.
- One commenter was concerned about the inconsistency between the level of analysis and level of implementation of the measure. The same commenter also raised concerns about the measure developer's decision not to include sociodemographic factors in the risk adjustment model.
- Committee Response:
  - The Committee agrees that the measure should be applied at the facility-level, as it is specified and tested. The Committee believes that linking claims and EHR data is an important advancement in quality measurement.
  - The Committee recognizes the potential for negative unintended consequences of admissions and readmissions measures and recommends careful monitoring of their implementation. Above all, the Committee agreed that use of these measures should be monitored to ensure they do not inadvertently reduce access to necessary care. The Committee noted the inverse relationship between mortality and readmission for heart failure and recognized the need for careful surveillance and balancing of these measures. The Committee also reiterated its concerns about the need to carefully balance implementation of measures addressing psychiatric readmissions to prevent the risk of higher suicidality.
    On the other hand, the Committee has noted the desire to understand a patient's need for any subsequent acute care after a hospitalization. In particular, the Committee recognized the need understand if patients are being seen in the Emergency Department after discharge or being placed in observation. The Committee recommends continued work to ensure that the use of readmissions measures does not result in unnecessary or avoidable use of the emergency department or observation status while ensuring that all patients have access to any necessary care. The Committee noted that a number of measures recommended for endorsement in this project could help to balance these concerns, in particular the measures addressing excess days in acute care and population-based admission measures.
  - The Committee endorsed this measure for facility-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.
    Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

    The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

    The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

2879 Hybrid Hospital-Wide Readmission Measure with Claims and Electronic Health Record Data

The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

- The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2880 Excess days in acute care (EDAC) after hospitalization for heart failure

Submission | Specifications

Description: This measure assesses days spent in acute care within 30 days of discharge from an inpatient hospitalization for heart failure to provide a patient-centered assessment of the post-discharge period. This measure is intended to capture the quality of care transitions provided to discharged patients hospitalized with heart failure by collectively measuring a set of adverse acute care outcomes that can occur post-discharge: emergency department (ED) visits, observation stays, and unplanned readmissions at any time during the 30 days post-discharge. In order to aggregate all three events, we measure each in terms of days. In 2016, CMS will begin annual reporting of the measure for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare, and are hospitalized in non-federal hospitals.

Numerator Statement: The outcome of the measure is a count of the number of days the patient spends in acute care within 30 days of discharge. We define days in acute care as days spent in an ED, admitted to an observation unit, or admitted as an unplanned readmission for any cause within 30 days from the date of discharge from the index heart failure hospitalization. Each ED treat-and-release visit is counted as one half-day (0.5 days). Observation stays are recorded in terms of hours and are rounded up to the nearest half-day. Each readmission day is counted as one full-day (1 day). We count all eligible outcomes occurring in the 30-day period, even if they are repeat occurrences.

Denominator Statement: The target population for this measure is Medicare FFS beneficiaries aged 65 years and older hospitalized at non-Federal acute care hospitals for heart failure.

The cohort includes admissions for patients discharged from the hospital with a principal discharge diagnosis of heart failure (see codes below in S.9) and with continuous 12 months Medicare enrollment prior to admission. The measure will be publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals.

Additional details are provided in S.9 Denominator Details.

Exclusions: The measure excludes index admissions for patients:

1. Without at least 30 days post-discharge enrollment in FFS Medicare.

2. Discharged against medical advice (AMA);

2880 Excess days in acute care (EDAC) after hospitalization for heart failure

3. Admitted within 30 days of a prior index discharge.

For 2016 public reporting, the measure will also exclude:

4. Admissions with a procedure code for left ventricular assist device (LVAD) implantation or heart transplantation either during the index admission or in the 12 months prior to the index admission. Patients with these procedures are a highly selected group of patients with different risk of the outcome. This exclusion will be added to the heart failure EDAC measure so that it remains fully harmonized with the CMS 30-day heart failure readmission measure. We did not exclude patients with LVAD or heart transplantation from the cohort of admissions used in the analyses for measure development and testing presented here.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-16; N-2**; 1b. Performance Gap: **H-7**; **M-11**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer cites that "the increasing use of ED visits and observation stays has raised concerns that current readmission measures do not capture the full range of unplanned acute care in the post-discharge period" (Vashi et al., 2013; Rising et al., 2012; Feng et al., 2012).
- Additionally, the developer notes that "observation stays can occur in many different parts of the hospital, including dedicated treatment rooms, the ED, or inpatient units. In particular, there is concern that high use of observation stays could in some cases replace readmissions, and that hospitals with high rates of observation stays in the post-discharge period may therefore have low readmission rates that do not accurately reflect the quality of care (Vashi et al., 2013)."
- Data provided by the developer cover a total of 575,672 discharges and show that heart failure readmission rates ranges from a minimum of -67 to a maximum of 196.
- The Standing Committee noted that the measure identifies a significant gap in performance with the 10th percentile at -29 days and the 90th percentile at 44.4 days. The Commmittee agreed that the measure met the NQF importance to measure and report criteria.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-0**; **M-17**; **L-1**; **I-0** 2b. Validity: **H-1**; **M-16**; **L-1**; **I-0**

<u>Rationale</u>:

- With regard to data element reliability, the developer notes that the measure has been developed to avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS. Additionally, the developer used the final risk-adjustment variables in the existing, NQF-endorsed measure of hospital-level risk-standardized readmission rates following AMI (NQF #0505).
- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method.
- For test-retest reliability, the developer calculated the EDAC for each hospital using first the development

2880 Excess days in acute care (EDAC) after hospitalization for heart failure

sample, then the validation sample. Thus, each hospital twice was measured twice, each time using an entirely distinct set of patients. The developer states that the extent to which the calculated measures of these two subsets agree is evidence that the measure is assessing an attribute of the hospital, not of the patients. As a metric of agreement, the developer calculated the intra-class correlation coefficient (ICC) as defined by ICC[2,1] by Shrout and Fleiss (1979) and assessed the values according to conventional standards (Landis and Koch, 1977).

- A total of 1,180,895 admissions were examined, with 590,448 in one sample and 590,447 in the other. The agreement between the two EDAC values for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.73.
- The developer demonstrated measure validity through prior validity testing done on their claims-based measures, through use of established measure development guidelines, and by systematic assessment of measure face validity by a Technical Expert Panel (TEP).
- The measure employs a hierarchical generalized linear model [HGLM]) that consists of two parts, a logit model and a truncated Poisson model. The two-part logit/Poisson model (often called a "hurdle" model) assumes that the outcome results from two related processes: an initial dichotomous event – that a patient has at least one acute care event – which is modeled as the logit of the probability of the event, and for patients with an event (those which clear the "hurdle"), the number of days, which is modeled as a Poisson process. The outcome, number of days, is a half-integer count variable (because ED visits count as 0.5 days).
- The developers also considered a number of variables related to sociodemographic status (SDS) for potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources.
- The developers state that both the patient-level and hospital-level dual eligible and race effects were significant in the logistic part of the HF EDAC model, but only the hospital-level effect was significant in the Poisson part of the model. This indicates that a) both the patient- and hospital-level dual eligible and race effects are associated with an increased risk of acute care but b) only the hospital-level effect is associated with the expected duration of that care. The developers note that if dual eligibility or race are used in the model to adjust for patient-level differences, then some of the differences between hospitals would also be adjusted for, potentially obscuring a signal of hospital quality.
- The developers state that given these findings and complex pathways that could explain any relationship between SDS and readmission, they did not incorporate SDS variables into the measure.
- For the logit model of zero versus non-zero days, which includes all patients in the cohort, the developers calculated the c-statistic.
    - C-statistic for logit part of model: 0.587
- For the Poisson model of non-zero days, which includes only patients with some acute care, the developers calculated the deviance R2. The deviance R2 is computed from the difference in the log-likelihoods between the final model and an empty model (no covariates) attributed to each observation, averaged over all observations.
    - Deviance R2 for truncated Poisson part of model: 0.026 (2.6%)
- Several Standing Committee members had concerns that this new methodology may cause confusion, since it is not the usual observed to expected ratio. Standing Committee members noted that this format for measure reporting may require education since it is not as consistent with the methods used in the past for other readmissions measures.
- The Standing Committee noted that unlike readmission rates, this measure captures a normalized number of days after hospitalization and may not be easily be compared across conditions.
- Standing Committee members noted that the empirical testing showed a Poisson correlation of 0.714 and the TEP agreement was around 92 percent, with 83 percent of the TEP in moderate or strong agreement. However, the Standing Committee had concerns about the c-statistic of 0.59, which is not very good.
- The Standing Committee agreed this measure met the reliability and validity criteria.

3. Feasibility: H-13; M-3; L-0; 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)

2880 Excess days in acute care (EDAC) after hospitalization for heart failure

Rationale:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure would be feasible to collect and implement.

4. Usability and Use: H-1; M-15; L-2; I-0

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

Rationale:

- This measure in not currently publicly reported, but was finalized for use in CMS' Hospital Inpatient Quality Report (IQR) program starting in FY 2018.

5. Related and Competing Measures

- This measure is related to NQF #0330: Hospital 30-day All-Cause RSRR Following Heart Failure Hospitalization. The developers note that both measures are harmonized.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-1**

Rationale

- The Standing Committee agreed that this measure is an important contribution to performance measurement as it captures the potiental unintended negative consequences of increased ED use and observation stays when measuring readmissions. Standing Committee members emphasized that the developers should communicate the differences between these measures and the readmissions measures so there is no confusion, since the reporting format is not as consistent with methods used in the past.

6. Public and Member Comment

- This measure received three comments. One comment expressed support for this measure to be recommended for endorsement.
- One commenter raised concerns about the potential unintended consequences of endorsing the measure, and the unknown number of truly preventable readmissions.
- One commenter was concerned about the inconsistency between the level of analysis and level of implementation of the measure. The same commenter also raised concerns about the measure developer's decision not to include sociodemographic factors in the risk adjustment model.
- Committee Response:
  - The Committee recognizes the potential for negative unintended consequences of admissions and readmissions measures and recommends careful monitoring of their implementation. Above all, the Committee agreed that use of these measures should be monitored to ensure they do not inadvertently reduce access to necessary care. The Committee noted the inverse relationship between mortality and readmission for heart failure and recognized the need for careful surveillance and balancing of these measures. The Committee also reiterated its concerns about the need to carefully balance implementation of measures addressing psychiatric readmissions to prevent the risk of higher suicidality.
    On the other hand, the Committee has noted the desire to understand a patient's need for any subsequent acute care after a hospitalization. In particular, the Committee recognized the need understand if patients are being seen in the Emergency Department after discharge or being placed in observation. The Committee recommends continued work to ensure that the use of readmissions measures does not result in unnecessary or avoidable use of the emergency department or observation status while ensuring that all patients have access to any necessary care. The Committee noted that a number of measures recommended for endorsement in this project could help to balance these concerns, in particular the measures addressing excess days in acute care and population-based admission measures.
  - The Committee endorsed this measure for facility-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder

2880 Excess days in acute care (EDAC) after hospitalization for heart failure
review. Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer. The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging. The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses. The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures. The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process. •
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2881 Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI)
Submission | Specifications
Description: This measure assesses days spent in acute care within 30 days of discharge from an inpatient hospitalization for acute myocardial infarction (AMI) to provide a patient-centered assessment of the post-discharge period. This measure is intended to capture the quality of care transitions provided to discharged patients hospitalized with AMI by collectively measuring a set of adverse acute care outcomes that can occur post-discharge: emergency department (ED) visits, observation stays, and unplanned readmissions at any time during the 30 days post-discharge. In order to aggregate all three events, we measure each in terms of days. In 2016, CMS will begin annual reporting of the measure for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare, and are hospitalized in non-federal hospitals. **Numerator Statement**: The outcome of the measure is a count of the number of days the patient spends in acute care within 30 days of discharge. We define days in acute care as days spent in an ED, admitted to an observation unit, or admitted as an unplanned readmission for any cause within 30 days from the date of discharge from the

2881 Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI)

index AMI hospitalization. Each ED treat-and-release visit is counted as one half-day (0.5 days). Observation stays are recorded in terms of hours and are rounded up to the nearest half-day. Each readmission day is counted as one full day (1 day). We count all eligible outcomes occurring in the 30-day period, even if they are repeat occurrences.

Denominator Statement: The target population for this measure is Medicare FFS beneficiaries aged 65 years and older hospitalized at non-federal acute care hospitals for AMI.

The cohort includes admissions for patients discharged from the hospital with a principal discharge diagnosis of AMI (see codes below in S.9) and with continuous 12 months Medicare enrollment prior to admission. The measure will be publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals.

Additional details are provided n S.9 Denominator Details.

Exclusions: The measure excludes index admissions for patients:

1. Without at least 30 days post-discharge enrollment in FFS Medicare;

2. Discharged against medical advice (AMA);

3. Admitted within 30 days of a prior index discharge;

4. Admitted and then discharged on the same day (because it is unlikely these are clinically significant AMIs).

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-18; N-0**; 1b. Performance Gap: **H-12**; **M-6**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer cites that "the increasing use of ED visits and observation stays has raised concerns that current readmission measures do not capture the full range of unplanned acute care in the post-discharge period" (Vashi et al., 2013; Rising et al., 2012; Feng et al., 2012).
- Additionally, the developer notes that "observation stays can occur in many different parts of the hospital, including dedicated treatment rooms, the ED, or inpatient units. In particular, there is concern that high use of observation stays could in some cases replace readmissions, and that hospitals with high rates of observation stays in the post-discharge period may therefore have low readmission rates that do not accurately reflect the quality of care (Vashi et al., 2013)."
- Data provided by the developer cover a total of 232,954 discharges and show that AMI readmission rates range from a minimum of -54 to a maximum of 170.
- Similar to NQF #2880, the Standing Committee agreed that the measure has a significant performance gap with the 10th percentile -23 days to the 90th percentile at 46 days among hospitals. The Standing Committee agreed that the measure is important to measure and report.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-0**; **M-18**; **L-0**; **I-0** 2b. Validity: **H-1**; **M-17**; **L-0**; **I-0**

<u>Rationale</u>:

- With regard to data element reliability, the developer notes that the measure has been developed to

2881 Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI)

avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS. Additionally, the developer used the final risk-adjustment variables in the existing, NQF-endorsed measure of hospital-level risk-standardized readmission rates following AMI (NQF #0505).

- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method.
- For test-retest reliability, the developer calculated the EDAC for each hospital using first the development sample, then the validation sample. Thus, each hospital twice was measured twice, each time using an entirely distinct set of patients. The developer states that the extent to which the calculated measures of these two subsets agree is evidence that the measure is assessing an attribute of the hospital, not of the patients. As a metric of agreement, the developer calculated the intra-class correlation coefficient (ICC) as defined by ICC[2,1] by Shrout and Fleiss (1979) and assessed the values according to conventional standards (Landis and Koch, 1977).
- A total of 496,716 admissions were examined, with 248,358 in each sample. The agreement between the two EDAC values for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.54.
- The developer demonstrated measure validity through prior validity testing done on their claims-based measures, through use of established measure development guidelines, and by systematic assessment of measure face validity by a Technical Expert Panel (TEP).
- The measure employs a hierarchical generalized linear model [HGLM]) that consists of two parts, a logit model and a truncated Poisson model. The two-part logit/Poisson model (often called a "hurdle" model) assumes that the outcome results from two related processes: an initial dichotomous event – that a patient has at least one acute care event – which is modeled as the logit of the probability of the event, and for patients with an event (those which clear the "hurdle"), the number of days, which is modeled as a Poisson process. The outcome, number of days, is a half-integer count variable (because ED visits count as 0.5 days).
- The developers considered a number of variables related to sociodemographic status (SDS) for potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources.
- The developers state that both the patient-level and hospital-level dual eligible and race effects were significant in the logistic part of the AMI EDAC model, but only the hospital-level effect was significant in the Poisson part of the model. This indicates that a) both the patient- and hospital-level dual eligible and race effects are associated with an increased risk of acute care but b) only the hospital-level effect is associated with the expected duration of that care. The developers note that if the dual eligible or race are used in the model to adjust for patient-level differences, then some of the differences between hospitals would also be adjusted for, potentially obscuring a signal of hospital quality.
- The developers state that given these findings and complex pathways that could explain any relationship between SDS and readmission, they did not incorporate SDS variables into the measure
- The Standing Committee had moderate certainty that the measure scores are reliable and valid with an intraclass correlation coefficient of 0.54, and a correlation with readmissions of 0.61.
- For the logit model of zero versus non-zero days, which includes all patients in the cohort, the developers calculated the c-statistic.
    - C-statistic for logit part of model: 0.60
- For the Poisson model of non-zero days, which includes only patients with some acute care, the developers calculated the deviance R2. The deviance R2 is computed from the difference in the log-likelihoods between the final model and an empty model (no covariates) attributed to each observation, averaged over all observations.
    - Deviance R2 for truncated Poisson part of model: 0.040 (4.0%)
- Standing Committee members expressed that the observed to predicted graph on this measure was better than the heart failure measure #2880.
- The Standing Committee agreed this measure met the reliability and validity criteria.

2881 Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI)

3. Feasibility: H-15; M-3; L-0; 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:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure would be feasible to collect and implement.

4. Usability and Use: H-4; M-14; L-0; I-0

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

Rationale:

- This measure in not currently publicly reported, but was finalized for use in CMS' Hospital Inpatient Quality Report (IQR) program starting in FY 2018.

5. Related and Competing Measures

- This measure is related to NQF #0505: Hospital 30-day all-cause risk-standardized readmission rate (RSRR) following acute myocardial infarction (AMI) hospitalization. The developers note that both measures are harmonized.

Standing Committee Recommendation for Endorsement: **Y-18**; **N-0**

Rationale

- The Standing Committee recognized the importance of reducing excess days in acute care due to acute myocardial infarction. The Standing Committee agreed that this measure is an important contribution to performance measurement as it captures the potential unintended negative consequences of increased ED use and observation stays when measuring readmissions. Standing Committee members emphasized that the developers should communicate the differences between these measures and the readmissions measures so there is no confusion, since the reporting format is not as consistent with the methods used in the past for readmissions measures.

6. Public and Member Comment

- This measure received four comments. One comment was in support of recommending the measure for endorsement.
- One commenter was concerned about the inconsistency between the level of analysis and level of implementation of the measure. The same commenter also raised concerns about the measure developer's decision not to include sociodemographic factors in the risk adjustment model.
- One commenter raised concerns about the level of reliability for this measure, saying the Intra-Class Correlation Coefficient (ICC) of 0.48 was low and an ICC of 0.60 should be the threshold.
- One commenter raised concerns about the intent of the measure and the utility of a measure that broadly defines acute care. The same commenter was concerned about the overlap of this measure and NQF #0505.
- Committee Response:
  - The Committee endorsed this measure for facility-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.

    Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

    The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current

2881 Excess days in acute care (EDAC) after hospitalization for acute myocardial infarction (AMI)

data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

- While the measure that was submitted to NQF has an Intra-Class Correlation Coefficient below 0.60, the Committee believes it represents an acceptable benchmark for reliability for measurement of excess days in acute care after hospitalization for AMI. The Committee concluded that developers' current approach to risk-adjustment and exclusions met the Scientific Acceptability criteria, and were satisfied with the measure's reliability.
- The Committee followed NQF's guidance on measure harmonization throughout the evaluation process. Prior to the in-person meeting, the Committee received materials regarding these competing measures, and held a separate call after the in-person meeting on September 1 to discuss harmonization issues and allow the developers to answer questions from Committee members. The Committee then voted via survey to recommend both measures. The Committee considered the added value and burden of recommending both measures and agreed that the differences in measure specifications added sufficient value to offset any potential negative impact.

- .

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2882 Excess days in acute care (EDAC) after hospitalization for pneumonia

Submission | Specifications

Description: This measure assesses days spent in acute care within 30 days of discharge from an inpatient hospitalization for pneumonia to provide a patient-centered assessment of the post-discharge period. This measure is intended to capture the quality of care transitions provided to discharged patients hospitalized with pneumonia by collectively measuring a set of adverse acute care outcomes that can occur post-discharge: emergency department (ED) visits, observation stays, and unplanned readmissions at any time during the 30 days

2882 Excess days in acute care (EDAC) after hospitalization for pneumonia

post-discharge. In order to aggregate all three events, we measure each in terms of days. In 2016, the Center for Medicare and Medicaid Services (CMS) will begin annual reporting of the measure for patients who are 65 years or older, are enrolled in fee-for-service (FFS) Medicare, and are hospitalized in non-federal hospitals.

Numerator Statement: The outcome of the measure is a count of the number of days the patient spends in acute care within 30 days of discharge. We define days in acute care as days spent in an ED, admitted to an observation unit, or admitted as an unplanned readmission for any cause within 30 days from the date of discharge from the index pneumonia hospitalization. Each ED treat-and-release visit is counted as one half-day (0.5 days). Observation stays are recorded in terms of hours and are rounded up to the nearest half-day. Each readmission day is counted as one full day (1 day). We count all eligible outcomes occurring in the 30-day period, even if they are repeat occurrences.

Denominator Statement: The target population for this measure is Medicare FFS beneficiaries aged 65 years and older hospitalized at non-Federal acute care hospitals for pneumonia.

The cohort includes admissions for patients discharged from the hospital with a principal discharge diagnosis of pneumonia (see codes below in S.9) and with continuous 12 months Medicare enrollment prior to admission. The measure will be publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals.

Additional details are provided n S.9 Denominator Details.

Exclusions: The measure excludes index admissions for patients:

1. Without at least 30 days post-discharge enrollment in FFS Medicare.

2. Discharged against medical advice (AMA);

3. Admitted within 30 days of a prior index discharge;

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" (Krumholz et al., 2006).

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-18; N-0**; 1b. Performance Gap: **H-13**; **M-4**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer cites that "the increasing use of ED visits and observation stays has raised concerns that current readmission measures do not capture the full range of unplanned acute care in the post-discharge period" (Vashi et al., 2013; Rising et al., 2012; Feng et al., 2012).
- Additionally, the developer notes that "observation stays can occur in many different parts of the hospital, including dedicated treatment rooms, the ED, or inpatient units. In particular, there is concern that high use of observation stays could in some cases replace readmissions, and that hospitals with high rates of observation stays in the post-discharge period may therefore have low readmission rates that do not accurately reflect the quality of care (Vashi et al., 2013)."
- Data provided by the developer cover a total of 495,130 discharges and show that pneumonia readmission rates ranged from a minimum of -67 to a maximum of 229.
- The Standing Committee agreed that the measure had fairly large performance gap that ranged from 67 days to 230 days and thus important to measure and report.

2882 Excess days in acute care (EDAC) after hospitalization for pneumonia

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-5**; **M-12**; **L-0**; **I-0** 2b. Validity: **H-0**; **M-17**; **L-1**; **I-0**

Rationale:

- With regard to data element reliability, the developer notes that the measure has been developed to avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS. Additionally, the developer used the final risk-adjustment variables in the current CMS 30-day pneumonia readmission measure.
- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method.
- For test-retest reliability, the developer calculated the EDAC for each hospital using first the development sample, then the validation sample. Thus, each hospital twice was measured twice, each time using an entirely distinct set of patients. The developer states that the extent to which the calculated measures of these two subsets agree is evidence that the measure is assessing an attribute of the hospital, not of the patients. As a metric of agreement, the developer calculated the intra-class correlation coefficient (ICC) as defined by ICC[2,1] by Shrout and Fleiss (1979) and assessed the values according to conventional standards (Landis and Koch, 1977).
- A total of 990,260 admissions were examined, with 495,130 in each sample. The agreement between the two EDAC values for each hospital (as measured by an intra-class correlation coefficient (ICC)) was 0.80.
- The developer demonstrated measure validity through prior validity testing done on their claims-based measures, through use of established measure development guidelines, and by systematic assessment of measure face validity by a Technical Expert Panel (TEP).
- The measure employs a hierarchical generalized linear model [HGLM]) that consists of two parts, a logit model and a truncated Poisson model. The two-part logit/Poisson model (often called a "hurdle" model) assumes that the outcome results from two related processes: an initial dichotomous event – that a patient has at least one acute care event – which is modeled as the logit of the probability of the event, and for patients with an event (those which clear the "hurdle"), the number of days, which is modeled as a Poisson process. The outcome, number of days, is a half-integer count variable (because ED visits count as 0.5 days).
- The developers considered a number of variables related to sociodemographic status (SDS) for potential inclusion in the risk-adjustment model. Candidate SDS variables were selected for examination based on a review of literature and national data sources.
- The developers state that both the patient-level and hospital-level dual eligible and race effects were significant in the logistic part of the pneumonia EDAC model, but only the hospital-level effect was significant in the Poisson part of the model. This indicates that a) both the patient- and hospital-level dual eligible and race effects are associated with an increased risk of acute care but b) only the hospital-level effect is associated with the expected duration of that care. The developers note that if the dual eligible or race are used in the model to adjust for patient-level differences, then some of the differences between hospitals would also be adjusted for, potentially obscuring a signal of hospital quality.
- The developers state that given these findings and complex pathways that could explain any relationship between SDS and readmission, they did not incorporate SDS variables into the measure.
- For the logit model of zero versus non-zero days, which includes all patients in the cohort, the developers calculated the c-statistic.
    - C-statistic for logit part of model: 0.616
- For the Poisson model of non-zero days, which includes only patients with some acute care, the developers calculated the deviance R2. The deviance R2 is computed from the difference in the log-likelihoods between the final model and an empty model (no covariates) attributed to each observation, averaged over all observations.

2882 Excess days in acute care (EDAC) after hospitalization for pneumonia

- Deviance R2 for truncated Poisson part of model: 0.034 (3.4%)

- The Standing Committee had moderate certainty that the measure scores are reliable and valid, with an intraclass correlation coefficient of 0.8, and a correlation with readmissions of 0.7. The face validity of the measure had a 91 percent agreement, of which 83 perfect were moderate or strong agreement.
- The Standing Committee agreed that this measure met the reliability and validity criteria and encouraged the developer to continue to test innovative approaches to improve the the prediction accuracy of this measure and others like it.

3. Feasibility: H-14; M-4; L-0; 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:

- This measure is collected through administrative claims data. The Standing Committee agreed the measure would be feasible to collect and implement.

4. Usability and Use: H-2; M-14; L-2; I-0

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

Rationale:

- This measure in not currently publicly reported, but may be used in one or more CMS programs, such as the IQR program.
- The Standing Committee agreed that this measure met the NQF usability and use criteria.

5. Related and Competing Measures

- This measure is related to NQF #0506: Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following pneumonia hospitalization. The developers note that both measures are harmonized.

Standing Committee Recommendation for Endorsement: **Y-18**; **N-0**

Rationale

- The Standing Committee recognized the importance of reducing excess days in acute care due to pneumonia. The Standing Committee agreed that this measure is an important contribution to performance measurement as it captures the potential unintended negative consequences of increased ED use and observation stays when measuring readmissions. Standing Committee members emphasized that the developers should communicate the differences between these measures and the readmissions measures so there is no confusion, since the reporting format is not as consistent with the methods used in the past for readmissions measures.

6. Public and Member Comment

- This measure received three comments. One comment submitted was in support of recommending the measure for endorsement.
- One commenter was concerned about the inconsistency between the level of analysis and level of implementation of the measure. The same commenter also raised concerns about the measure developer's decision not to include sociodemographic factors in the risk adjustment model.
- One commenter raised concerns about the intent of the measure and the utility of a measure that broadly defines acute care. The same commenter was also concerned about the overlap of this measure and NQF #0506.
- Committee Response:
  - The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.
  - Consideration of sociodemographic factors in risk adjustment models is a critical issue in

2882 Excess days in acute care (EDAC) after hospitalization for pneumonia

measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

- The Committee followed NQF's guidance on measure harmonization throughout the evaluation process. Prior to the in-person meeting, the Committee received materials regarding these competing measures, and held a separate call after the in-person meeting on September 1 to discuss harmonization issues and allow the developers to answer questions from Committee members. The Committee then voted via survey to recommend both measures. The Committee considered the added value and burden of recommending both measures and agreed that the differences in measure specifications added sufficient value to offset any potential negative impact.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-; N-

9. Appeals

2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure

Submission | Specifications

Description: Rate of risk-standardized acute, unplanned hospital admissions among Medicare Fee-for-Service (FFS) patients 65 years and older with heart failure

Numerator Statement: The outcome measured for each patient is the number of acute, unplanned admissions per 100 person-years at risk for admission. Persons are considered at risk for admission if they are alive, enrolled in FFS Medicare, and not currently admitted. (See S.6, Numerator Details, for more information.)

2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure

Denominator Statement: The target population is ambulatory Medicare FFS patients aged 65 years and older with a diagnosis of heart failure.

Exclusions: The measure excludes:

1. Patients without continuous enrollment in Medicare Part A for the duration of the measurement period (or until death).

Rationale: We exclude these patients to ensure full data availability for outcome assessment (Part A during the measurement year).

2. Patients with left ventricular assist devices (LVADs).

Rationale: We exclude these patients because while they have a high risk of admission, they are low in prevalence and are clustered among a few ACOs.

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: We use a two-level hierarchical negative binomial model to estimate risk-standardized acute, unplanned admissions per person-year at risk for admission. This approach accounts for the clustering of patients within ACOs and variation in sample size.

Level of Analysis: Integrated Delivery System

Setting of Care: Ambulatory Care : Clinician Office/Clinic, Other

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-18; N-0**; 1b. Performance Gap: **H-11**; **M-8**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer provided data from ACO performance score using the 2012 Medicare Full Sample which showed the crude US national Medicare FFS rate of acute, unplanned admissions per person-year among patients with heart failure was 85.5 per 100 person-years.
    - Among ACOs, the mean RSAAR for calendar year 2012 was 81.9 per 100 person-years (standard deviation = 11.6). The median RSAAR was 81.5 per 100 person-years (interquartile range [IQR] 73.6 to 88.8). The minimum RSAAR score was 53.7 per 100 person-years; the 5th percentile was 64.6 per 100 person-years; the 95th percentile was 101.7 per 100 person-years; and maximum score was 120.7 per 100 person-years.
    - They observed that 61 ACOs (53.5%) had RSAARs that were 'no different than the national rate' (of all Medicare FFS beneficiaries with heart failure). An additional 37 ACOs (32.5%) had 'better than the national rate' RSAAR scores and 16 (14.0%) were 'worse than the national rate, which signaled a gap in performance to the Standing Committee.
- The Standing Committee agreed that this measure fills an important gap and there is evidence of the relationship between clinical interventions and the ability to prevent hospitalizations. The Standing Committee noted that this measure will be helpful to accoutable care organizations (ACOs) and they attempt to improve quality and better understand their total costs but did express concerns that the measure could be challenging to use in a quality initiative program when the interventions to improve take time to establish and ACOs enter the program at different times.
- The Standing Committee suggested that future directions for measurement in this area could assess ED use, observation stays, and skilled nursing facility admissions.
- The Standing Committee agreed the measure met the evidence criterion.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-2**; **M-18**; **L-0**; **I-0** 2b. Validity: **M-14**; **L-6**; **I-0**

2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure

Rationale:

- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method. This is generally considered an appropriate method of testing reliability.
    - The 2012 full Medicare sample was divided into two subsets of patients randomly. The developer calculated the measure score of all ACOs for each of the two subsets of patients. Each ACO was measured twice, but each measurement was made using distinct sets of measures. The interclass correlation coefficient (ICC) for the two subsets of patients was 0.81, which can be interpreted as excellent correlation, and thus reliable
- The Standing Committee raised concerns about the impact of sample size on reliability and questioned if there was a need for a minimum number of cases, particularly if the measure were to be applied to sample ACOs.
- The Standing Committee noted that this measure is calculated using fee-for-service claims and questioned how the transition to alternative payment models could impact this measure.
- The Standing Committee recommended that the developer continue to refine this measure to expand the population to patients under 65 to capture understudied populations and to promote public-private sector alignment.
- The developers provided a conceptual framework that was used to develop the risk adjustment model for this measure. This conceptual framework included 4 contextual domains that influence ACO performance including, physical environment, community resources, patient resources, and patient behavioral/personal preferences.
- The measure included demographic factors, and clinical risk factors present at the start of the measurement period.
- The measure developers reviewed 189 diagnosis groups included in the hierarchical condition category (HCC), and calculated the prevalence of each CC in the year preceding the measurement period. After examining the bi-variate analysis, the developers reduced the list to 22 candidate variables including age.
- The measure developers did not adjust for contextual factors that impact admissions; however, they did provide data demonstrating that including SDS adjustment did not make a meaningful difference to the measure score of the ACOs. The spearman correlation coefficient that estimated the difference in performance with and without SDS adjustment was 0.990. Thus, the results demonstrate that adjustment had little effect on the measure score.
- To assess the overall performance of their risk-adjustment model, the developers computed two summary statistics, including:
    - *Risk model discrimination statistics* (the model's ability to explain how successful the fit is in explaining the variation of the data. In this case, the r-squared value was 0.123. In other words, the model was able to explain 12.3% of the total deviance.
    - Overfitting indices (model calibration) [presented as (γ0, γ1)]:
        - The developer states that if the γ0 in the validation samples are substantially far from zero and the γ1 is substantially far from one, there is potential evidence of over-fitting. The calibration value of close to 0 at one end and close to 1 to the other end indicates good calibration of the model.
            - 2012 Development Sample (Index): (**0,1**)
            - 2012 Validation Sample: **(-0.0020, 1.0002)**
- Ultimately the Standing Committee agreed that this measure was reliable.
- The developer tested the validity of the measure using three different methods:
    - Validity of the claims-based measures. The developer argues that other NQF endorsed mortality and readmission measures have been validated by comparing the claims to the medical records data elements. It is unclear if the risk adjustment validation approach that the developer cites is sufficiently similar to this measure and for this level of analysis and ambulatory patients.
    - The developer also notes that this measure has been validated by using established measure development guidelines. While an important step for measure development, this method of

2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure

validity testing has generally not been considered sufficient for demonstrating measure validity.
    - Finally, the measure developer completed a systemic face validity assessment of this measure with 8 experts agreeing that this measure was a valid indicator of health care quality
- While the Standing Committee ultimately supported the developer's decision not to adjust for SDS factors that some of those factors did show a significant effect.
- .

3. Feasibility: H-12; M-7; L-0; 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:

- All measure data elements are in defined fields in electronic claims and routinely generated or collected by and used by healthcare personnel during the provision of care, coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims).
- The Standing Committee agreed the measure was feasible.

4. Usability and Use: H-5; M-14; L-0; I-0

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

Rationale:

- The measure is not currently used for public reporting or in an accountability program. However, this measure was included by CMS in the November 2014 Physician Fee Schedule final rule, and finalized adding the measure to the Medicare Shared Savings Program (MSSP) quality measure set. The measure is planned for pay-for-performance in the MSSP for 2017 reporting period.
- The Standing Committee agreed that this measure was useful but raised concerns that it may overlap with how CMS is using NQF #0277: Heart Failure Admission Rate (PQI 8) in the MSSP program.

5. Related and Competing Measures

- The Standing Committee raised concerns that that this measure may compete with NQF #0277 : Heart Failure Admission Rate (PQI 8), which calculates admissions with a principal diagnosis of heart failure per 100,000 population, ages 18 years and older and excludes cardiac procedure admissions, obstetric admissions, and transfers from other institutions. Measure #0277 and Measure #2886 both calculate the admissions of patients with heart failure.
- Measure #0277 measures those who are aged 18 years and older, and Measure #2886 only measures those aged 65 years and older.
- The Standing Committee will review these issues during a follow up call.

Standing Committee Recommendation for Endorsement: **Y-19**; **N-1**

Rationale

- The Standing Committee recognized the importance of reducing unplanned hospital admissions among Medicare Fee-for-Service (FFS) patients 65 years and older with heart failure and recommended the measure for endorsement.

6. Public and Member Comment

- This measure received three comments. One comment was in support of recommending this measure for endorsement.
- One commenter was concerned about the inconsistency between the level of analysis and level of implementation of the measure. The same commenter also raised concerns about the measure developer's decision not to include sociodemographic factors in the risk adjustment model.
- One comment raised concerns that the risk adjustment model did not adequately address concerns for sociodemographic factors specifically for ACOs.
- Committee Response:

2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure

- The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.
  Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

  The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

  The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

  The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

  The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

- Developer Response:
  - The goal of risk adjustment is to ensure that the measure is fair and reflects differences in quality, not case mix. Thus, we adjusted for factors that affect patients' risk of admission, not quality that ACOs can and should influence.
    We did not adjust for non-clinical contextual factors since it is within the mission of ACOs to partner with their communities to improve population health.

    We conducted several analyses to demonstrate that ACOs in different contextual environments have the capacity to do well on our measure. In the publicly available methodology report, we show heterogeneity in performance among ACOs with the most and fewest number of patients who were dual eligible and of low socioeconomic status.

    As part of this work, we are considering ways to further characterize the diverse contextual environments and patients ACOs serve. We agree it would be informative to understand how additional factors -- such as the physical environment, health behaviors, and social and economic environments – influence risk-adjusted admission rates. However, this is a new area for quality assessment and methods are evolving.

2886 Risk-Standardized Acute Admission Rates for Patients with Heart Failure
In summary, while we appreciate the concern that ACOs caring for higher volumes of patients from poorer and less resourced communities face challenges, some of these ACOs do well. Our conceptual model and data support not including SDS-related factors in the risk-adjustment model.
7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0
8. Board of Directors Vote: Y-X; N-X
9. Appeals

2887 Risk-Standardized Acute Admission Rates for Patients with Diabetes
Submission | Specifications
Description: Rate of risk-standardized acute, unplanned hospital admissions among Medicare fee-for-service (FFS) patients 65 years and older with diabetes **Numerator Statement**: The outcome measured for each patient is the number of acute, unplanned admissions per 100 person-years at risk for admission. Persons are considered at risk for admission if they are alive, enrolled in FFS Medicare, and not currently admitted. (See S.6, Numerator Details, for more information.) **Denominator Statement**: The target population is ambulatory Medicare FFS patients aged 65 years and older with a diagnosis of diabetes. **Exclusions**: The measure excludes: 1. Patients without continuous enrollment in Medicare Part A for the duration of the measurement period (or until death). Rationale: We exclude these patients to ensure full data availability for outcome assessment (Part A during the measurement year). **Adjustment/Stratification**: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: We use a two-level hierarchical negative binomial model to estimate risk-standardized acute, unplanned admissions per person-year at risk for admission. This approach accounts for the clustering of patients within ACOs and variation in sample size. **Level of Analysis:** Integrated Delivery System **Setting of Care:** Ambulatory Care : Clinician Office/Clinic, Other **Type of Measure**: Outcome **Data Source**: Administrative claims **Measure Steward**: Centers for Medicare & Medicaid Services (CMS)
STANDING COMMITTEE MEETING [06/08/2016-06/09/2016] **1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>** (1a. Evidence, 1b. Performance Gap) 1a. Evidence: **Y-19; N-0**; 1b. Performance Gap: **H-7**; **M-13**; **L-0**; **I-0** <u>Rationale</u>: • The developer provided data from ACO performance score using the 2012 Medicare Full Sample which showed the mean risk-standardized acute admission rate (RSAAR) among ACOs for year 2012 is 39.6, median is 39.1.   ○ They observed that 51 ACOs (44.7%) had RSAARs that were 'no different than the national rate' and 45 ACOs (39.5%) had RSAAR scores 'better than the national rate,' and 18 ACOs (15.8%) were 'worse than the national rate', which signaled a gap in performance to the Standing Committee. • The Standing Committee agreed that this measure fills an important gap and there is evidence of the relationship between clinical interventions and the ability to prevent hospitalizations. The Standing Committee also noted that the measure shows evidence of disparities in care. • The Standing Committee agreed the measure met the evidence criterion.

2887 Risk-Standardized Acute Admission Rates for Patients with Diabetes

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-2**; **M-17**; **L-0**; **I-0** 2b. Validity: **M-17**; **L-3**; **I-0**

<u>Rationale</u>:

- Datasets used for testing included Medicare Parts A and B claims, the denominator file, the Medicare Provider Analysis and Review (MedPAR) file, and the American Community Survey to derive the AHRQ SES index.
- With regard to data element reliability, the developer notes that the measure has been developed to avoid the use of claims data elements that are thought to be coded inconsistently across hospitals or providers, instead using fields that are consequential for payment and which are audited by CMS.
    - Summarizing the results of this analysis, the developer notes that the mean age and frequency of risk-adjustment variables was similar among the two samples of 2012 data suggesting that the data elements are reliable across the samples.
- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method.
    - The 2012 full Medicare sample was divided into two subsets of patients randomly. The developer calculated the measure score of all ACOs for each of the two subsets of patients. Each ACO was measured twice, but each measurement was made using distinct sets of measures. The interclass correlation coefficient (ICC) for the two subsets of patients was 0.889, which can be interpreted as excellent correlation, and thus reliable.
- The developer tested the validity of the measure using three different methods:
    - Validity of the claims-based measures. The developer argues that other NQF endorsed mortality and readmission measures have been validated by comparing the claims to the medical records data elements. It is unclear if the risk adjustment validation approach that the developer cites is sufficiently similar to this measure and for this level of analysis and ambulatory patients.
    - The developer also notes that this measure has been validated by using established measure development guidelines. While an important step for measure development, this method of validity testing has generally not been considered sufficient for demonstrating measure validity.
    - Finally, the measure developer completed a systemic face validity assessment of this measure with 9 experts and two patients agreeing that this measure was a valid indicator of health care quality.
- The developers provided a conceptual framework that was used to develop the risk adjustment model for this measure. This conceptual framework included 4 contextual domains that influence ACO performance including, physical environment, community resources, patient resources, and patient behavioral/personal preferences.
- The measure included demographic factors, and clinical risk factors present at the start of the measurement period.
- The measure developers reviewed 189 diagnosis groups included in the hierarchical condition category (HCC), and calculated the prevalence of each CC in the year preceding the measurement period. After examining the bi-variate analysis, the developers reduced the list to 22 candidate variables including age.
- The measure developers did not adjust for contextual factors that impact admissions; however, they did provide data demonstrating that including SDS adjustment did not make a meaningful difference to the measure score of the ACOs. The spearman correlation coefficient that estimated the difference in performance with and without SDS adjustment was 0.981. Thus, the results demonstrate that adjustment had little effect on the measure score.
- To assess the overall performance of their risk-adjustment model, the developers computed two summary statistics, including:
    - *Risk model discrimination statistics* (the model's ability to explain how successful the fit is in explaining the variation of the data. In this case, the r-squared value was 0.218. In other words,

2887 Risk-Standardized Acute Admission Rates for Patients with Diabetes

the model was able to explain 21.8% of the total deviance.

- Overfitting indices (model calibration) [presented as (γ0, γ1)]:
  - The developer states that if the γ0 in the validation samples are substantially far from zero and the γ1 is substantially far from one, there is potential evidence of over-fitting. The calibration value of close to 0 at one end and close to 1 to the other end indicates good calibration of the model.
    - 2012 Development Sample (Index): (**0,1**)
    - 2012 Validation Sample: **(0.0017, 1.0031)**

- The Standing Committee noted that the developed decided not to include SDS factors despite some change in model performance due to concerns about disparities and variations in performance.
- The Standing Committee raised questions about the classification of wound debridement as a planned admission and therefore excluded from the measure. Ultimately the Standing Committee agreed with the developer's algorithm for exclusions.
- The Standing Committee noted the impact that self-selection bias could have on the results of this measure. Higher performing providers may be opting into forming ACOs leading to challenges comparing scores to the national average.

3. Feasibility: H-15; M-5; L-0; 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:

- All measure data elements are in defined fields in electronic claims and routinely generated or collected by and used by healthcare personnel during the provision of care, coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims).
- The Standing Committee agreed the measure was feasible.

4. Usability and Use: H-5; M-14; L-1; I-0

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

Rationale:

- The measure is not currently used for public reporting or in an accountability program. However, this measure was included by CMS in the November 2014 Physician Fee Schedule final rule, and finalized adding the measure to the Medicare Shared Savings Program (MSSP) quality measure set. The measure is planned for pay-for-performance in the MSSP for 2017 reporting period.
- Given the importance of managing diabetes in the ambulatory setting, the Standing Committee recommended that the developer explore ways to expand the admissions included in the measure. The Standing Committee noted that not all planned care represents a good outcome for the patient. Additionally the Standing Committee stressed the need not provide a disincentive to necessary acute care.

5. Related and Competing Measures

- This measure may compete with NQF #0272: Diabetes Short-Term Complications Admission Rate (PQI 01), which calculates admissions for a principal diagnosis of diabetes with short-term complications (ketoacidosis, hyperosmolarity, or coma) per 100,000 population, ages 18 years and older and excludes obstetric admissions and transfers from other institutions.
- This measure may compete with NQF #0274: Diabetes Long-Term Complications Admission Rate (PQI 03), which calculates Admissions for a principal diagnosis of diabetes with long-term complications (renal, eye, neurological, circulatory, or complications not otherwise specified) per 100,000 population, ages 18 years and older and excludes obstetric admissions and transfers from other institutions.
- This measure may compete with NQF #0638: Uncontrolled Diabetes Admission Rate (PQI 14), which calculates the admissions for a principal diagnosis of diabetes without mention of short-term (ketoacidosis, hyperosmolarity, or coma) or long-term (renal, eye, neurological, circulatory, or other unspecified) complications per 100,000 population, ages 18 years and older.

2887 Risk-Standardized Acute Admission Rates for Patients with Diabetes

- All three of these related measures are also outcome measures and also measure admissions rates for patients with diabetes. Measures #0272, 0274, and 0638 measure those aged 18 years and older but Measure #2887 is only for those aged 65 years and older.
- Measures #0272, 0274, and 0638 are all in the hospital setting while Measure #2887 is in the ambulatory care setting.
- The Standing Committee will review these issues during a follow up call

Standing Committee Recommendation for Endorsement: **Y-18**; **N-2**

Rationale

- The Standing Committee recognized the importance of reducing unplanned hospital admissions among Medicare Fee-for-Service (FFS) patients 65 years and older with diabetes and recommended the measure for endorsement.

6. Public and Member Comment

- This measure received three comments. One comment was in support of recommending this measure for endorsement.
- One commenter was concerned about the inconsistency between the level of analysis and level of implementation of the measure. The same commenter also raised concerns about the measure developer's decision not to include sociodemographic factors in the risk adjustment model.
- One comment raised concerns that the risk adjustment model did not adequately address concerns for sociodemographic factors specifically for ACOs.
- Committee Response:
  - The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.

    Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

    The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

    The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

    The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

    The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues

2887 Risk-Standardized Acute Admission Rates for Patients with Diabetes

to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

- Developer Response:
  - The goal of risk adjustment is to ensure that the measure is fair and reflects differences in quality, not case mix. Thus, we adjusted for factors that affect patients' risk of admission, not quality that ACOs can and should influence.
    We did not adjust for non-clinical contextual factors since it is within the mission of ACOs to partner with their communities to improve population health.

    We conducted several analyses to demonstrate that ACOs in different contextual environments have the capacity to do well on our measure. In the publicly available methodology report, we show heterogeneity in performance among ACOs with the most and fewest number of patients who were dual eligible and of low socioeconomic status.

    As part of this work, we are considering ways to further characterize the diverse contextual environments and patients ACOs serve. We agree it would be informative to understand how additional factors -- such as the physical environment, health behaviors, and social and economic environments – influence risk-adjusted admission rates. However, this is a new area for quality assessment and methods are evolving.
  - In summary, while we appreciate the concern that ACOs caring for higher volumes of patients from poorer and less resourced communities face challenges, some of these ACOs do well. Our conceptual model and data support not including SDS-related factors in the risk-adjustment model.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-X; N-X

9. Appeals

2888 Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions

Submission | Specifications

Description: Rate of risk-standardized acute, unplanned hospital admissions among Medicare fee-for-service (FFS) patients 65 years and older with multiple chronic conditions (MCCs)

Numerator Statement: The outcome measured for each patient is the number of acute, unplanned admissions per 100 person-years at risk for admission. Persons are considered at risk for admission if they are alive, enrolled in FFS Medicare, and not currently admitted. (See S.6, Numerator Details, for more information.)

Denominator Statement: Our target population is Medicare FFS patients aged 65 years and older whose combinations of chronic conditions put them at high risk of admission and whose admission rates could be lowered through better care. The National Quality Forum's (NQF's) "Multiple Chronic Conditions Measurement Framework," which defines patients with multiple chronic conditions as people "having two or more concurrent chronic conditions that…. act together to significantly increase the complexity of management, and affect functional roles and health outcomes, compromise life expectancy, or hinder self-management [1]."

Operationally, the measure cohort includes patients with diagnoses in two or more of eight chronic disease groups:

1. Acute myocardial infarction (AMI)
2. Alzheimer's disease and related disorders or senile dementia
3. Atrial fibrillation
4. Chronic kidney disease (CKD)

2888 Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions

5. Chronic obstructive pulmonary disease (COPD) and asthma

6. Depression

7. Heart failure

8. Stroke and transient ischemic attack (TIA)

This approach captures approximately 25% of Medicare FFS beneficiaries aged 65 years and older with at least one chronic condition (about 5 million patients in 2012).

Citations:

1. National Quality Forum (NQF). Multiple Chronic Conditions Measurement Framework. 2012; http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=71227

Exclusions: The measure excludes:

1. Patients without continuous enrollment in Medicare Part A for the duration of the measurement period (or until death).

Rationale: We exclude these patients to ensure full data availability for outcome assessment (Part A during the measurement year).

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: We use a two-level hierarchical negative binomial model to estimate risk-standardized acute, unplanned admissions per 100 person-years at risk for admission. This approach accounts for the clustering of patients within ACOs and variation in sample size.

Level of Analysis: Integrated Delivery System

Setting of Care: Ambulatory Care : Clinician Office/Clinic

Type of Measure: Outcome

Data Source: Administrative claims

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

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

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

1a. Evidence: **Y-20; N-0**; 1b. Performance Gap: **H-4**; **M-16**; **L-0**; **I-0**

<u>Rationale</u>:

- The developer noted improvements in access to care, supporting self-care in the home, coordinating care across providers, and integrating social work, nursing, and medical services all have the potential to improve admission rates for patients with multiple chronic conditions.
- Using data from the 2012 Medicare Full Sample with 4,937,344 patients, that was composed of 239,551 patients in 114 ACOs, and compared with the 71.9 admissions (per 100 person-years) - the US national Medicare FFS rate of acute, unplanned admissions among patients with MCCs, they found that:
  - The mean risk-standardized acute admission rate (RSAAR) among ACOs for year 2012 was 69.3, median was 68.5.
  - They observed that 45 ACOs (39.5%) had RSAARs that were 'no different than the national rate' and 22 ACOs (19.3%) had RSAAR scores 'worse than the national rate,' and 47 ACOs (41.2%) were 'better than the national rate", which signaled a gap in performance to the Standing Committee.
- The Standing Committee noted the need to for measures assessing multiple chronic conditions. The Standing Committee felt this measure could be an important first step to assessing the impact of frailty on readmissions.
- The Standing Committee felt there was a performance gap and that there were interventions an ACO could perform to improve performance.
- The Standing Committee agreed that this measure met the evidence criteria.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2888 Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions

2a. Reliability: **H-3**; **M-17**; **L-0**; **I-0** 2b. Validity: **M-16**; **L-4**; **I-0**

Rationale:

- The developer's approach to assessing score-level reliability was to consider the extent to which assessments of a hospital using different but randomly-selected subsets of patients produce similar measures of hospital performance. The developers refer to this as a "test-retest" approach; it may also be called a "split-half" method. This is generally considered an appropriate method of testing reliability.
    - The 2012 full Medicare sample was divided into two subsets of patients randomly. The developer calculated the measure score of all ACOs for each of the two subsets of patients. Each ACO was measured twice, but each measurement was made using distinct sets of measures. The interclass correlation coefficient (ICC) for the two subsets of patients was 0.84, which can be interpreted as excellent correlation, and thus reliable.
- This measure estimates the predicted number of admissions given the Accountable Care Organization's (ACO's) case mix, sample size, and actual admission rate. The outcome for this measure is the number of acute, unplanned admissions per 100 person-years at risk for admission. The outcome includes inpatient admissions to an acute care hospital for any cause during the measurement year, unless an admission is identified as "planned."
- The developer tested the validity of the measure using three different methods:
    - Validity of the claims-based measures. The developer argues that other NQF endorsed mortality and readmission measures have been validated by comparing the claims to the medical records data elements. It is unclear if the risk adjustment validation approach that the developer cites is sufficiently similar to this measure and for this level of analysis and ambulatory patients.
    - The developer also notes that this measure has been validated by using established measure development guidelines. While an important step for measure development, this method of validity testing has generally not been considered sufficient for demonstrating measure validity.
    - Finally, the measure developer completed a systemic face validity assessment of this measure with 9 experts and two patients agreeing that this measure was a valid indicator of health care quality.
- The developers provided a conceptual framework that was used to develop the risk adjustment model for this measure. This conceptual framework included 4 contextual domains that influence ACO performance including, physical environment, community resources, patient resources, and patient behavioral/personal preferences.
- The measure included demographic factors, and clinical risk factors present at the start of the measurement period.
- The measure developers reviewed 189 diagnosis groups included in the hierarchical condition category (HCC), and calculated the prevalence of each CC in the year preceding the measurement period. After examining the bi-variate analysis, the developers reduced the list to 46 candidate variables including age.
- The measure developers did not adjust for contextual factors that impact admissions; however, they did provide data demonstrating that including SDS adjustment did not make a meaningful difference to the measure score of the ACOs. The spearman correlation coefficient that estimated the difference in performance with and without SDS adjustment was 0.992. Thus, the results demonstrate that adjustment had little effect on the measure score.
- To assess the overall performance of their risk-adjustment model, the developers computed two summary statistics, including:
    - Risk model discrimination statistics (the model's ability to explain how successful the fit is in explaining the variation of the data. In this case, the r-squared value was 0.123. In other words, the model was able to explain 12.3% of the total deviance.
    - Overfitting indices (model calibration) [presented as (γ0, γ1)]:
        - The developer states that if the γ0 in the validation samples are substantially far from zero and the γ1 is substantially far from one, there is potential evidence of over-fitting. The calibration value of close to 0 at one end and close to 1 to the other end indicates good calibration of the model.
            - 2012 Development Sample (Index): (0,1)

2888 Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions

- 2012 Validation Sample: (-0.0015, 1.0011)
- Given the complexity of this measure, the Standing Committee raised concerns about converting the data from ICD-9 to ICD-10; ultimately the Standing Committee agreed the measure was valid.
- The Standing Committee agreed this measure met the scientific acceptability criteria.

3. Feasibility: H-12; M-8; L-0; 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:

- All measure data elements are in defined fields in electronic claims and routinely generated or collected by and used by healthcare personnel during the provision of care, coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims).
- The Standing Committee agreed the measure was feasible.

4. Usability and Use: H-4; M-16; L-0; I-0

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

Rationale:

- The measure is not currently used for public reporting or in an accountability program. However, this measure was included by CMS in the November 2014 Physician Fee Schedule final rule, and finalized adding the measure to the Medicare Shared Savings Program (MSSP) quality measure set. The measure is planned for pay-for-performance in the MSSP for 2017 reporting period.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-20**; **N-0**

Rationale

- The Standing Committee recognized the importance of reducing unplanned hospital admissions among Medicare Fee-for-Service (FFS) patients 65 years and older with MCCs and recommended the measure for endorsement.

6. Public and Member Comment

- This measure received four comments. Two comments were in support of recommending this measure for endorsement.
- One commenter was concerned about the inconsistency between the level of analysis and level of implementation of the measure. The same commenter also raised concerns about the measure developer's decision not to include sociodemographic factors in the risk adjustment model.
- One comment raised concerns that the risk adjustment model did not adequately address concerns for sociodemographic factors specifically for ACOs.
- Committee Response:
  - The Committee endorsed this measure for hospital-level analysis based on the testing results submitted for review. The Committee agrees that this measure should not be used for individual or group practices unless updated testing and specifications are provided to the Standing Committee to support endorsement for that use case. The Committee encourages the measure developer to bring additional testing results for alternative use cases to NQF for multistakeholder review.

    Consideration of sociodemographic factors in risk adjustment models is a critical issue in measurement science. The Committee takes the concerns raised by the commenters seriously. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer.

    The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and socio-demographic risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers

2888 Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions

to test the conceptual and empirical relationship between SDS factors and readmissions. The Committee's deliberations on the need for SDS adjustment were challenging.

The Committee noted particular limitations for measures that were conditionally endorsed based on the need for review under the NQF trial period for SDS adjustment. The committee acknowledged that measure developers were not required to address social determinants in the original analyses required for NQF review and endorsement, which contributed to the relative lack of data to ensure robust assessment of the impact of SDS in many of the post-hoc analyses.

The Committee reiterated that their focus was on the adjustments the developer was able to put forward at this time given the data currently available. While the adjustments put forward for these measures at this time did not reach a threshold of significance the Committee was comfortable with the Committee recognizes that risk adjustment for SDS factors is a rapidly progressing area and that more work is needed to appreciate the effects of social risk, understand the most relevant patient-and community level risk factors, collect data on these risk factors, and determine the best methods to incorporate these risk factors into performance measures.

The Committee stressed the high risk of unintended consequences related to adjustment of these measures for SDS factors and the need to reevaluate these measures as the field continues to move forwards. The Committee recognized the need to ensure facilities serving vulnerable populations are not penalized unfairly while at the same time balancing concerns about worsening healthcare disparities. The Committee looks forward to continued deliberations on these issues and to reexamining these measures as better data emerges. The Committee recommends a reassessment of the availability of SDS variables and a reexamination of these measures through the NQF annual update process.

- Developer Response:
  - The goal of risk adjustment is to ensure that the measure is fair and reflects differences in quality, not case mix. Thus, we adjusted for factors that affect patients' risk of admission, not quality that ACOs can and should influence.
    We did not adjust for non-clinical contextual factors since it is within the mission of ACOs to partner with their communities to improve population health.

    We conducted several analyses to demonstrate that ACOs in different contextual environments have the capacity to do well on our measure. In the publicly available methodology report, we show heterogeneity in performance among ACOs with the most and fewest number of patients who were dual eligible and of low socioeconomic status.

    As part of this work, we are considering ways to further characterize the diverse contextual environments and patients ACOs serve. We agree it would be informative to understand how additional factors -- such as the physical environment, health behaviors, and social and economic environments – influence risk-adjusted admission rates. However, this is a new area for quality assessment and methods are evolving.

    In summary, while we appreciate the concern that ACOs caring for higher volumes of patients from poorer and less resourced communities face challenges, some of these ACOs do well. Our conceptual model and data support not including SDS-related factors in the risk-adjustment model.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-12; N-0

8. Board of Directors Vote: Y-X; N-X

9. Appeals

Measure Not Recommended

2884 30-Day Unplanned Readmissions for Cancer Patients

Submission

Description: 30-Day Unplanned Readmissions for Cancer Patients is a cancer-specific measure. It provides the rate at which all adult cancer patients (= 18 years old), regardless of payer type, have an unplanned re-hospitalization within 30 days of an index admission. The readmission is defined as a subsequent inpatient admission to the reporting facility, which occurs within 30 days of the discharge date of an eligible index admission.

Numerator Statement: This outcome measure demonstrates the rate at which adult cancer patients (=18 years old at the index admission) are readmitted to a PPS-exempt Cancer Hospital (PCH) within 30 days of discharge from an index admission at the same PCH. The numerator includes all eligible patients with a readmission to a PCH within 30 days of the discharge date from an index admission with an admission status of urgent or emergency

Denominator Statement: All adult inpatient admissions with a diagnosis of malignant cancer at PCHs over the defined measurement period. The outcome measure examines the rate of unplanned readmissions within 30 days of discharge of this population.

Exclusions: The following patients are excluded from the denominator population: 1) patients transferred to another acute care facility during the index admission; 2) having missing or incomplete data; 3) admitted to an inpatient hospice bed; and, 4) discharged Against Medical Device (AMA).

Adjustment/Stratification: Risk Adjustment Type: Statistical risk model; Risk Adjustment Methodology: A logistic regression was applied, using the following risk factors: 1) age less than 40; 2) discharge to hospice; 3) length of stay greater than 3 days; 4) low socioeconomic status; 5) multiple comorbidities; 6) solid tumor; and, 7) Surgical MS-DRG.

Level of Analysis: Facility

Setting of Care: Hospital/Acute Care Facility

Type of Measure: Outcome

Data Source: Administrative claims

Measure Steward: Seattle

STANDING COMMITTEE MEETING [06/08/2016-06/09/2016]

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

(1a. Evidence: 1b. Performance Gap, 1c. High Priority)

1a. Evidence: **Y-17; N-0;** 1b. Performance Gap: **H-1**; **M-16**; **L-0**; **I-1**

Rationale:

- Cancer is the second leading cause of death in the United States. Approximately 1.7 million Americans are diagnosed with cancer each year but there is no measure to assess readmission rates for this disease. Cancer patients are also currently excluded from all-cause readmission rates such as NQF #1789.
- Unadjusted readmission rates to dedicated cancer facilities range from 14.5 percent to 15.8 percent.
- For many patients readmission may be preventable and should be addressed to lower costs and improve patient outcomes. Readmissions may be prevented by ensuring adequate treatment during the index hospitalization and post-discharge.
- The Standing Committee recommended that the developers separate out payer class as a marker of socioeconomic challenges. In particular the Standing Committee raised concerns about the unique challenges Medicaid patients face when seeking treatment for cancer and recommended that they not be categorized with patients who are opting to pay for treatment out of pocket.
- The Standing Committee also suggested the developer consider ways to track readmissions to other facilities and to consider a longer time window.

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

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

2884 30-Day Unplanned Readmissions for Cancer Patients

2a. Reliability: **H-0**; **M-5**; **L-13**; **I-1**

Rationale:

- The developer has assessed reliability at the data element level. The reliability of the measure was testing by comparing the level of agreement with the planned/unplanned indicator based on the sample chart review. A Kappa score was calculated for the overall agreement of the two measures and the facility-level agreement.
- Inter-rater reliability analyses (Kappa) were performed to determine consistency between Planned/Unplanned readmission type and inclusion in the measure numerator for individual participating facilities. Kappa scores ranged from 0.080 to 1.000 with asymptotic standard error ranging from 0.000 to 0.113.
- The developer notes that a moderate level of agreement (0.772) resulted when Kappa scores across the ten participating facilities were averaged. However, while seven out of the ten participating facilities have Kappa scores above 0.800, three centers had scores ranging from 0.080 to 0.690. Variation in applied definitions of "planned" and/or "unplanned" readmissions is one explanation for the widespread Kappa scores. A second source of variation may be the internal facility's guidelines for determining the type of admission. Third, some variation may be due to numerator exclusion criteria (i.e., admissions with a primary diagnosis of chemotherapy or radiation therapy encounter or progression of disease).
- The Standing Committee raised concerns about the performance of this outlier and that it may be challenging to implement this measure broadly.
- The Standing Committee also noted that the measure only tracks readmissions to the same facility. However, a patient could be readmitted to a different facility. The Standing Committee had concerns that a hospital's location and a patient's ability to seek care at a different facility could impact the reliability of the measure. Variability in rates could be driven by the healthcare market in a given location rather than facility quality.
- The measure did not pass reliability.

[i] "Health Policy Brief: Medicare Hospital Readmissions Reduction Program," Health Affairs, November 12, 2013.
[ii] Medicare Payment Advisory Committee (MEDPAC). Report to the Congress: Medicare and the Health Care Delivery System, DC: MedPAC; 2013. Available at http://medpac.gov/documents/Jun13_EntireReport.pdf