NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 0277 **NQF Project:** Cardiovascular Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Congestive Heart Failure Admission Rate (PQI 8)
De.2 **Brief description of measure***:* Percent of county population with an admissions for CHF.
1.1-2 **Type of Measure:** Outcome De.3 **If included in a composite or paired with another measure, please identify composite or paired measure** Prevention Quality Indicator (PQI) composite
De.4 **National Priority Partners Priority Area:** Population health, Safety De.5 **IOM Quality Domain**: Effectiveness De.6 **Consumer Care Need:** Staying healthy

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** A.3 **Measure Steward Agreement:** Government entity and in the public domain - no agreement necessary A.4 **Measure Steward Agreement attached:**	**A** **Y**☒ **N**☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** **Y**☒ **N**☐

C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement	**C** Y☒ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** Yes, fully developed and tested D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☒ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☒ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s): RWinkler	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. *Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**: Population health	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** High resource use, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Billings et al. found that low-income ZIP codes in New York City had 4.6 times more CHF hospitalizations per capita than high-income ZIP codes. Millman et al. reported that low-income ZIP codes had 6.1 times more CHF hospitalizations per capita than high-income ZIP codes.65 Based on empirical results, areas with high rates of CHF also tend to have high rates of admission for other ACSCs. 1a.4 **Citations for Evidence of High Impact:** Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report. Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington DC: National Academy Press.	**1a** C☐ P☐ M☐ N☐
1b. Opportunity for Improvement 1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** Congestive heart failure is a PQI that would be of most interest to comprehensive health care delivery systems. This indicator is measured with high precision, and most of the observed variance reflects true differences across areas. Risk adjustment for age and sex appears to affect the areas with the highest and lowest raw rates. Areas with high rates may wish to examine the clinical characteristics of their patients to check for a more complex case mix. Patient age, clinical measures such as heart function, and other management issues may affect admission rates.	**1b** C☐ P☐ M☐ N☐

Comment [KP1]: 1a. The measure focus addresses:
- a specific national health goal/priority identified by NQF's National Priorities Partners; OR
- a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

Comment [KP2]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

As the causes for admissions may include poor quality care, lack of patient compliance, or problems accessing care, areas may wish to review CHF patient records to identify precipitating causes and potential targets for intervention.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
Adjusted per 100,000 rates by patient and hospital characteristics, 2007

Mean	Standard error	Location	P-value: Relative to Northeast
402.605	22.318	Northeast	1.000
446.773	21.686	Midwest	0.156
474.166	17.900	South	0.012
293.022	11.579	West	0.000

1b.3 Citations for data on performance gap:
See the following report for a complete treatment of the methodology: "Methods: Applying AHRQ Quality Indicators to Healthcare Cost and Utilization Project (HCUP) Data for the National Healthcare Quality Report" [URL: http://hcupnet.ahrq.gov/QI%20Methods.pdf?JS=Y]

1b.4 Summary of Data on disparities by population group:
Adjusted per 100,000 rates by patient characteristics, 2007

Estimate	Standard error	Age: for conditions affecting any age
38.527	1.828	18-44
298.394	10.627	45-64
1912.391	43.139	65 and over

Estimate	Standard error	Age: for conditions affecting elderly
835.456	22.964	65-69
1243.6	30.172	70-74
1845.486	43.594	75-79
2841.152	69.354	80-84
4453.902	114.115	85 and over

Estimate	Standard error	Gender
474.842	11.383	Male
370.707	8.504	Female

Estimate	Standard error	Median income of patient´s ZIP code
561.781	25.3	First quartile (lowest income)
420.838	16.952	Second quartile
361.98	14.697	Third quartile
319.623	20.016	Fourth quartile (highest income)

Estimate	Standard error	Location of patient residence (NCHS)
442.037	34.923	Large central metropolitan
413.407	31.738	Large fringe metropolitan
380.89	36.494	Medium metropolitan
398.905	45.931	Small metropolitan
417.946	23.022	Micropolitan
430.314	20.094	Not metropolitan or micropolitan

1b.5 Citations for data on Disparities:
See the following report for a complete treatment of the methodology: "Methods: Applying AHRQ Quality Indicators to Healthcare Cost and Utilization Project (HCUP) Data for the National Healthcare Quality Report" [URL: http://hcupnet.ahrq.gov/QI%20Methods.pdf?JS=Y]

1c. Outcome or Evidence to Support Measure Focus — **1c** C☐

Comment [k3]: 1 Examples of data on opportunity for improvement include, but are not limited to: prior studies, epidemiologic data, measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.

Comment [k4]: 1c. The measure focus is:
- an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;

OR
- if an intermediate outcome, process, structure, etc., there is **evidence** that supports the specific measure focus as follows:
  - Intermediate outcome – evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
  - Process – evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
  - Structure – evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
  - Patient experience – evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
  - Access – evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
  - Efficiency – demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** Congestive heart failure (CHF) can be controlled in an outpatient setting for the most part. If area rates for CHF are high even after risk adjustment and stratification, the quality of preventive services in that region are held to be insufficient in preparing CHF patients to manage their condition.

P☐ M☐ N☐

1c.2-3. Type of Evidence: Evidence-based guideline, Expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**
Congestive heart failure is a PQI that would be of most interest to comprehensive health care delivery systems. This indicator is measured with high precision, and most of the observed variance reflects true differences across areas.
Risk adjustment for age and sex appears to affect the areas with the highest and lowest raw rates. Areas with high rates may wish to examine the clinical characteristics of their patients to check for a more complex case mix. Patient age, clinical measures such as heart function, and other management issues may affect admission rates.
As the causes for admissions may include poor quality care, lack of patient compliance, or problems accessing care, areas may wish to review CHF patient records to identify precipitating causes and potential targets for intervention.

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**
RATING: 14 Testing, rating, and review were conducted by the project team. A full report on the literature review and empirical evaluation can be found in Refinement of the HCUP Quality Indicators by the UCSF-Stanford EPC, Detailed coding information for each QI is provided in the document Prevention Quality Indicators Technical Specifications. Rating of performance on empirical evaluations, ranged from 0 to 26. The scores were intended as a guide for summarizing the performance of each indicator on four empirical tests of precision (signal variance, area-level share, signal ratio, and R-squared) and five tests of minimum bias (rank correlation, top and bottom decile movement, absolute change, and change over two deciles), as described in the previous section.

1c.6 Method for rating evidence: The project team conducted extensive empirical testing of all potential indicators using the 1995-97 HCUP State Inpatient Databases (SID) and Nationwide Inpatient Sample (NIS) to determine precision, bias, and construct validity. The 1997 SID contains uniform data on inpatient stays in community hospitals for 22 States covering approximately 60% of all U.S. hospital discharges. The NIS is designed to approximate a 20% of U.S. community hospitals and includes all stays in the sampled hospitals. Each year of the NIS contains between 6 million and 7 million records from about 1,000 hospitals. The NIS combines a subset of the SID data, hospital-level variables, and hospital and discharge weights for producing national estimates. The project team conducted tests to examine three things: precision, bias, and construct validity.
Precision. The first step in the analysis involved precision tests to determine the reliability of the indicator for distinguishing real differences in provider performance. For indicators that may be used for quality improvement, it is important to know with what precision, or surety, a measure can be attributed to an actual construct rather than random variation.
For each indicator, the variance can be broken down into three components: variation within a provider (actual differences in performance due to differing patient characteristics), variation among providers (actual differences in performance among providers), and random variation. An ideal indicator would have a substantial amount of the variance explained by between-provider variance, possibly resulting from differences in quality of care, and a minimum amount of random variation. The project team performed four tests of precision to estimate the magnitude of between-provider variance on each indicator:
• Signal standard deviation was used to measure the extent to which performance of the QI varies systematically across hospitals or areas.
• Provider/area variation share was used to calculate the percentage of signal (or true) variance relative to the total variance of the QI.
• Signal-to-noise ratio was used to measure the percentage of the apparent variation in QIs across providers that is truly related to systematic differences across providers and not random variations (noise) from year to year.
• In-sample R-squared was used to identify the incremental benefit of applying multivariate signal extraction

Comment [k5]: 4 Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multi-step process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending immunization are necessary steps, they are not sufficient to achieve the desired impact on health status – patients must be vaccinated to achieve immunity. This does not preclude consideration of measures of preventive screening interventions where there is a strong link with desired outcomes (e.g., mammography) or measures for multiple care processes that affect a single outcome.

Comment [k6]: 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., USPSTF grading system http://www.ahrq.gov/clinic/uspstf07/methods/benefit.htm). If the USPSTF grading system was not used, the grading system is explained including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies and the best type of evidence depends upon the question being studied (e.g., randomized controlled trials appropriate for studying drug efficacy are not well suited for complex system changes). When qualitative studies are used, appropriate qualitative research criteria are used to judge the strength of the evidence.

methods for identifying additional signal on top of the signal-to-noise ratio.
In general, random variation is most problematic when there are relatively few observations per provider, when adverse outcome rates are relatively low, and when providers have little control over patient outcomes or variation in important processes of care is minimal. If a large number of patient factors that are difficult to observe influence whether or not a patient has an adverse outcome, it may be difficult to separate the "quality signal" from the surrounding noise. Two signal extraction techniques were applied to improve the precision of an indicator:
• Univariate methods were used to estimate the "true" quality signal of an indicator based on information from the specific indicator and 1 year of data.
• Multivariate signal extraction (MSX) methods were used to estimate the "true" quality signal based on information from a set of indicators and multiple years of data. In most cases, MSX methods extracted additional signal, which provided much more precise estimates of true hospital or area quality.
Bias. To determine the sensitivity of potential QIs to bias from differences in patient severity, unadjusted performance measures for specific hospitals were compared with performance measures that had been adjusted for age and gender. All of the PQIs and some of the Inpatient Quality Indicators (IQIs) could only be risk-adjusted for age and sex. The 3M™ APR-DRG System Version 12 with Severity of Illness and Risk of Mortality subclasses was used for risk adjustment of the utilization indicators and the in-hospital mortality indicators, respectively. Five empirical tests were performed to investigate the degree of bias in an indicator:
• Rank correlation coefficient of the area or hospital with (and without) risk adjustment—gives the overall impact of risk adjustment on relative provider or area performance.
• Average absolute value of change relative to mean—highlights the amount of absolute change in performance, without reference to other providers' performance.
• Percentage of highly ranked hospitals that remain in high decile—reports the percentage of hospitals or areas that are in the highest deciles without risk adjustment that remain there after risk adjustment is performed.
• Percentage of lowly ranked hospitals that remain in low decile—reports the percentage of hospitals or areas that are in the lowest deciles without risk adjustment that remain there after risk adjustment is performed.
• Percentage that change more than two deciles—identifies the percentage of hospitals whose relative rank changes by a substantial percentage (more than 20%) with and without risk adjustment.
Construct validity. Construct validity analyses provided information regarding the relatedness or independence of the indicators. If quality indicators do indeed measure quality, then two measures of the same construct would be expected to yield similar results. The team used factor analysis to reveal underlying patterns among large numbers of variables—in this case, to measure the degree of relatedness between indicators. In addition, they analyzed correlation matrices for indicators.

1c.7 **Summary of Controversy/Contradictory Evidence:** See the following for a complete treatment of the topic: http://www.qualityindicators.ahrq.gov/downloads/pqi/pqi_guide_v31.pdf
Note: The Literature Review Findings column summarizes evidence specific to each potential concern on the link between the PQIs and quality of care, as described in step 3 above. A question mark (?) indicates that the concern is theoretical or suggested, but no specific evidence was found in the literature. A check mark indicates that the concern has been demonstrated in the literature.

1c.8 **Citations for Evidence (***other than guidelines***):**
http://www.qualityindicators.ahrq.gov/downloads/pqi/pqi_guide_v31.pdf

1c.9 **Quote the Specific guideline recommendation (***including guideline number and/or page number***):**
Not Applicable.

1c.10 **Clinical Practice Guideline Citation:** Not Applicable.
1c.11 **National Guideline Clearinghouse or other URL:** Not Applicable.

1c.12 **Rating of strength of recommendation (***also provide narrative description of the rating and by whom***):**
Not Applicable.

1c.13 **Method for rating strength of recommendation** (*If different from USPSTF system, also describe rating and how it relates to USPSTF)***:**
Not Applicable.

Comment [k7]: USPSTF grading system **http://www.ahrq.gov/clinic/uspstf/grades.htm**: **A -** The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B -** The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C -** The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D -** The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I -** The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

1c.14 Rationale for using this guideline over others: Not Applicable.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, **met?** **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2 If yes, provide web page URL:**	
2a. Precisely Specified	
2a.1 Numerator Statement (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome***):** All discharges of age 18 years and older with ICD-9-CM principal diagnosis code for CHF. **2a.2 Numerator Time Window (***The time period in which cases are eligible for inclusion in the numerator***):** Time period is user defined. Users of the measure typically use a 12 month time period. **2a.3 Numerator Details (***All information required to collect/calculate the numerator, including all codes, logic, and definitions***):** All discharges of age 18 years and older with ICD-9-CM principal diagnosis code for CHF. Include ICD-9-CM diagnosis codes: 39891 RHEUMATIC HEART FAILURE 4280 CONGESTIVE HEART FAILURE 4281 LEFT HEART FAILURE 42820 SYSTOLIC HRT FAILURE NOS OCT02- 42821 AC SYSTOLIC HRT FAILURE OCT02- 42822 CHR SYSTOLIC HRT FAILURE OCT02- 42823 AC ON CHR SYST HRT FAIL OCT02- 42830 DIASTOLC HRT FAILURE NOS OCT02- 42831 AC DIASTOLIC HRT FAILURE OCT02- 42832 CHR DIASTOLIC HRT FAIL OCT02- 42833 AC ON CHR DIAST HRT FAIL OCT02- 42840 SYST/DIAST HRT FAIL NOS OCT02- 42841 AC SYST/DIASTOL HRT FAIL OCT02-	**2a-specs** **C☐** **P☐** **M☐** **N☐**

Comment [KP8]: 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF's Health Information Technology Expert Panel (HITEP) .

42842
CHR SYST/DIASTL HRT FAIL OCT02-
42843
AC/CHR SYST/DIA HRT FAIL OCT02-
4289
HEART FAILURE NOS
Include ICD-9-CM diagnosis codes ONLY for discharges before 2002Q3 (ending September 30, 2002):
40201
MAL HYPERT HRT DIS W CHF
40211
BENIGN HYP HRT DIS W CHF
40291
HYPERTEN HEART DIS W CHF
40401
MAL HYPER HRT/REN W CHF
40403
MAL HYP HRT/REN W CHF/RF
40411
BEN HYPER HRT/REN W CHF
40413
BEN HYP HRT/REN W CHF/RF
40491
HYPER HRT/REN NOS W CHF
40493
HYP HT/REN NOS W CHF/RF

Exclude cases:
• transfer from a hospital (different facility)
• transfer from a skilled Nursing Facility (SNF) or Intermediate Care Facility (ICF)
• transfer from another health care facility
• MDC 14 (pregnancy, childbirth, and puerperium)
• with a cardiac procedure code
ICD-9-CM Cardiac procedure codes
0050
IMPL CRT PACEMAKER SYS OCT02-
0051
IMPL CRT DEFIBRILLAT OCT02-
0052
IMP/REP LEAD LF VEN SYS OCT02-
0053
IMP/REP CRT PACEMKR GEN OCT02-
0054
IMP/REP CRT DEFIB GENAT OCT02-
0056
INS/REP IMPL SENSOR LEAD OCT06-
0057
IMP/REP SUBCUE CARD DEV OCT06-
0066
PTCA OCT06-
1751
IMPLANTATION OF RECHARGEABLE CARDIAC CONTRACTILITY MODULATION [CCM], TOTAL SYSTEM OCT09-
1752
IMPLANTATION OR REPLACEMENT OF CARDIAC CONTRACTILITY MODULATION [CCM] RECHARGEABLE PULSE GENERATOR ONLY OCT09-
3500
CLOSED VALVOTOMY NOS
3501
CLOSED AORTIC VALVOTOMY

3502 CLOSED MITRAL VALVOTOMY 3503 CLOSED PULMON VALVOTOMY 3504 CLOSED TRICUSP VALVOTOMY 3510 OPEN VALVULOPLASTY NOS 3511 OPN AORTIC VALVULOPLASTY 3512 OPN MITRAL VALVULOPLASTY 3513 OPN PULMON VALVULOPLASTY 3514 OPN TRICUS VALVULOPLASTY 3520 REPLACE HEART VALVE NOS 3521 REPLACE AORT VALV-TISSUE 3522 REPLACE AORTIC VALVE NEC 3523 REPLACE MITR VALV-TISSUE 3524 REPLACE MITRAL VALVE NEC 3525 REPLACE PULM VALV-TISSUE 3526 REPLACE PULMON VALVE NEC 3527 REPLACE TRIC VALV-TISSUE 3528 REPLACE TRICUSP VALV NEC 3531 PAPILLARY MUSCLE OPS 3532 CHORDAE TENDINEAE OPS 3533 ANNULOPLASTY 3534 INFUNDIBULECTOMY 3535 TRABECUL CARNEAE CORD OP 3539 TISS ADJ TO VALV OPS NEC 3541 ENLARGE EXISTING SEP DEF 3542 CREATE SEPTAL DEFECT 3550 PROSTH REP HRT SEPTA NOS 3551 PROS REP ATRIAL DEF-OPN 3552 PROS REPAIR ATRIA DEF-CL 3553 PROST REPAIR VENTRIC DEF	

3554
PROS REP ENDOCAR CUSHION
3555
PROS REP VENTRC DEF-CLOS OCT06-
3560
GRFT REPAIR HRT SEPT NOS
3561
GRAFT REPAIR ATRIAL DEF
3562
GRAFT REPAIR VENTRIC DEF
3563
GRFT REP ENDOCAR CUSHION
3570
HEART SEPTA REPAIR NOS
3571
ATRIA SEPTA DEF REP NEC
3572
VENTR SEPTA DEF REP NEC
3573
ENDOCAR CUSHION REP NEC
3581
TOT REPAIR TETRAL FALLOT
3582
TOTAL REPAIR OF TAPVC
3583
TOT REP TRUNCUS ARTERIOS
3584
TOT COR TRANSPOS GRT VES
3591
INTERAT VEN RETRN TRANSP
3592
CONDUIT RT VENT-PUL ART
3593
CONDUIT LEFT VENTR-AORTA
3594
CONDUIT ARTIUM-PULM ART
3595
HEART REPAIR REVISION
3596
PERC HEART VALVULOPLASTY
3598
OTHER HEART SEPTA OPS
3599
OTHER HEART VALVE OPS
3601
PTCA-1 VESSEL W/O AGENT
3602
PTCA-1 VESSEL WITH AGNT
3603
OPEN CORONRY ANGIOPLASTY
3604
INTRCORONRY THROMB INFUS
3605
PTCA-MULTIPLE VESSEL
3606
INSERT OF COR ART STENT OCT95-
3607
INS DRUG-ELUT CORONRY ST OCT02-

3609
REM OF COR ART OBSTR NEC
3610
AORTOCORONARY BYPASS NOS
3611
AORTOCOR BYPAS-1 COR ART
3612
AORTOCOR BYPAS-2 COR ART
3613
AORTOCOR BYPAS-3 COR ART
3614
AORTCOR BYPAS-4+ COR ART
3615
1 INT MAM-COR ART BYPASS
3616
2 INT MAM-COR ART BYPASS
3617
ABD-CORON ART BYPASS OCT96-
3619
HRT REVAS BYPS ANAS NEC
362
ARTERIAL IMPLANT REVASC
363
OTH HEART REVASCULAR
3631
OPEN CHEST TRANS REVASC
3632
OTH TRANSMYO REVASCULAR
3633
ENDO TRANSMYO REVASCULAR OCT06-
3634
PERC TRANSMYO REVASCULAR OCT06-
3639
OTH HEART REVASULAR
3691
CORON VESS ANEURYSM REP
3699
HEART VESSLE OP NEC
3731
PERICARDIECTOMY
3732
HEART ANEURYSM EXCISION
3733
EXC/DEST HRT LESION OPEN
3734
EXC/DEST HRT LES OTHER
3735
PARTIAL VENTRICULECTOMY
3736
EXCISION OR DESTRUCTION OF LEFT ATRIAL APPENDAGE (LAA) OCT08-
3741
IMPLANT PROSTH CARD SUPPORT DEV OCT06
375
HEART TRANSPLANTATION (NOT VALID AFTER OCT 03)
3751
HEART TRANPLANTATION OCT03-
3752
IMPLANT TOT REP HRT SYS OCT03-

3753
REPL/REP THORAC UNIT HRT OCT03-
3754
REPL/REP OTH TOT HRT SYS OCT03-
3755
REMOVAL OF INTERNAL BIVENTRICULAR HEART REPLACEMENT SYSTEM OCT08-
3760
IMPLANTATION OR INSERTION OF BIVENTRICULAR EXTERNAL HEART ASSIST SYSTEM OCT08-
3761
IMPLANT OF PULSATION BALLOON
3762
INSERTION OF NON-IMPLANTABLE HEART ASSIST SYSTEM
3763
REPAIR OF HEART ASSIST SYSTEM
3764
REMOVAL OF HEART ASSIST SYSTEM
3765
IMPLANT OF EXTERNAL HEART ASSIST SYSTEM
3766
INSERTION OF IMPLANTABLE HEART ASSIST SYSTEM
3770
INT INSERT PACEMAK LEAD
3771
INT INSERT LEAD IN VENT
3772
INT INSERT LEAD ATRI-VENT
3773
INT INSER LEAD IN ATRIUM
3774
INT OR REPL LEAD EPICAR
AHRQ Quality Indicators Web Site: http://www.qualityindicators.ahrq.gov
Prevention Quality Indicators Technical Specifications Version 4.2- 2010
PQI #8 Congestive Heart Failure (CHF) Admission Rate Page 3
3775
REVISION OF LEAD
3776
REPL TV ATRI-VENT LEAD
3777
REMOVAL OF LEAD W/O REPL
3778
INSER TEAM PACEMAKER SYS
3779
REVIS OR RELOCATE POCKET
3780
INT OR REPL PERM PACEMKR
3781
INT INSERT 1-CHAM, NON
3782
INT INSERT 1-CHAM, RATE
3783
INT INSERT DUAL-CHAM DEV
3785
REPL PACEM W 1-CHAM, NON
3786
REPL PACEM 1-CHAM, RATE
3787
REPL PACEM W DUAL-CHAM
3789

REVISE OR REMOVE PACEMAK
3794
IMPLT/REPL CARDDEFIB TOT
3795
IMPLT CARDIODEFIB LEADS
3796
IMPLT CARDIODEFIB GENATR
3797
REPL CARDIODEFIB LEADS
3798
REPL CARDIODEFIB GENRATR

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured***):**
Population in Metro Area or county, age 18 years and older.

2a.5 Target population gender: Female, Male
2a.6 Target population age range: 18 and older

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator***):**
Time period is user defined. Users of the measure typically use a 12 month time period.

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
Population in Metro Area or county, age 18 years and older.

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population***):** none

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions***):**
Not applicable

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***):**
Observed rates may be stratified by gender, age (5-year age groups), race / ethnicity

2a.12-13 Risk Adjustment Type: Risk adjustment method widely or commercially available

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method***):**
The predicted value for each case is computed using a logistic regression model and covariates for gender and age in years (in 5-year age groups). The reference population used in the model is the universe of discharges for states that participate in the HCUP State Inpatient Databases (SID) for the year 2007 (updated annually), a database consisting of 43 states and approximately 30 million adult discharges. The expected rate is computed as the sum of the predicted value for each case divided by the number of cases for the unit of analysis of interest (i.e., county, state, and region). The risk adjusted rate is computed using indirect standardization as the observed rate divided by the expected rate, multiplied by the reference population rate

2a.15-17 Detailed risk model available Web page URL or attachment: URL None
http://qualityindicators.ahrq.gov/downloads/pqi/PQI_Risk_Adjustment_Tables_(Version_4_2).pdf

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**
Each indicator is expressed as a rate, is defined as outcome of interest / population at risk or numerator / denominator. The AHRQ Quality Indicators (AHRQ QI) software performs five steps to produce the rates. 1) Discharge-level data is used to mark inpatient records containing the outcome of interest and 2) the population at risk. For provider indicators, the population at risk is also derived from hospital discharge records; for area indicators, the population at risk is derived from U.S. Census data. 3) Calculate observed

Comment [k9]: 11 Risk factors that influence outcomes should not be specified as exclusions.
12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

rates. Using output from steps 1 and 2, rates are calculated for user-specified combinations of stratifiers. 4) Calculate expected rates. Regression coefficients from a reference population database are applied to the discharge records and aggregated to the provider or area level. 5) Calculate risk-adjusted rate. Use the indirect standardization to account for case-mix. 6) Calculate smoothed rate. A Univariate shrinkage factor is applied to the risk-adjusted rates. The shrinkage estimate reflects a reliability adjustment unique to each indicator. Full information on calculation algorithms and specifications can be found at http://qualityindicators.ahrq.gov/PQI_download.htm

2a.22 **Describe the method for discriminating performance** *(e.g., significance testing)***:**
Significance testing is not prescribed by the software. Users may calculate a confidence interval for the risk-adjusted rates and a posterior probability interval for the smoothed rates at a 95% or 99% level. Users may define the relevant benchmark and the methods of discriminating performance according to their application.

2a.23 **Sampling (Survey) Methodology** *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
Not applicable

2a.24 **Data Source (***Check the source(s) for which the measure is specified and tested)*
Electronic administrative data/claims

2a.25 **Data source/data collection instrument (***Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**
The data source is hospital discharge data such as the HCUP State Inpatient Databases (SID) or equivalent using UB-04 coding standards. The data collection instrument is public-use AHRQ QI software available in SAS or Windows versions.

2a.26-28 **Data source/data collection instrument reference web page URL or attachment:** URL None
http://www.qualityindicators.ahrq.gov/software.htm

2a.29-31 **Data dictionary/code table web page URL or attachment:** URL None
http://www.qualityindicators.ahrq.gov/downloads/winqi/AHRQ_QI_Windows_Software_Documentation_V41a.pdf

2a.32-35 **Level of Measurement/Analysis** (*Check the level(s) for which the measure is specified and tested)*
Population: states, Population: counties or cities

2a.36-37 **Care Settings (***Check the setting(s) for which the measure is specified and tested)*
Ambulatory Care: Office

2a.38-41 **Clinical Services** (*Healthcare services being measured, check all that apply*)
Clinicians: Physicians (MD/DO)

TESTING/ANALYSIS

2b. Reliability testing

2b.1 **Data/sample** *(description of data/sample and size)***:** AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges

2b.2 **Analytic Method** *(type of reliability & rationale, method for testing)***:**
Expert panels and empirical analysis

2b.3 **Testing Results** *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)***:**
Relatively precise estimates of admission rates for CHF can be obtained, although random variation may be important for small hospitals and rural areas. Based on empirical evidence, this indicator is very precise, with a raw area level rate of 521.0 per 100,000 population and a standard deviation of 286.5. The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is very high, at 93.0%, indicating that the observed differences in age-sex adjusted rates very likely represent true differences across areas.

2b
C☐
P☐
M☐
N☐

Comment [KP10]: 2b. Reliability testing demonstrates the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

Comment [k11]: 8 Examples of reliability testing include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing may address the data items or final measure score.

2c. Validity testing

2c.1 **Data/sample** *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges

2c.2 **Analytic Method** *(type of validity & rationale, method for testing)*:
Expert panels and empirical analysis

2c.3 **Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
Billings et al. found that low-income ZIP codes in New York City had 4.6 times more CHF hospitalizations per capita than high-income ZIP codes.64 Millman et al. reported that low-income ZIP codes had 6.1 times more CHF hospitalizations per capita than high-income ZIP codes.65
Based on empirical results, areas with high rates of CHF also tend to have high rates of admission for other ACSCs.

2c
C☐ P☐ M☐ N☐

2d. Exclusions Justified

2d.1 **Summary of Evidence supporting exclusion(s)**:
Exclusions remove cases where the outcome of interest is less likely to be preventable or with no or very low risk

2d.2 **Citations for Evidence:**
Refinement of the HCUP Quality Indicators (Technical Review), May 2001
http://qualityindicators.ahrq.gov/downloads/technical/qi_technical_review.zip

2d.3 **Data/sample** *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges

2d.4 **Analytic Method** *(type analysis & rationale)*:
Expert panel and descriptive analyses stratified by exclusion categories

2d.5 **Testing Results** *(e.g., frequency, variability, sensitivity analyses)*:
Refinement of the HCUP Quality Indicators (Technical Review), May 2001
http://qualityindicators.ahrq.gov/downloads/technical/qi_technical_review.zip

2d
C☐ P☐ M☐ N☐ NA☐

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 **Data/sample** *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges

2e.2 **Analytic Method** *(type of risk adjustment, analysis, & rationale)*:
Risk-adjustment models use a standard set of categories based on readily available classification systems for demographics, severity of illness and comorbidities. Within each category, covariates are initially selected based on a minimum of 30 cases in the outcome of interest. Then a stepwise regression process on a development sample is used to select a parsimonious set of covariates where p<.05. Model is then tested on a validation sample

2e.3 **Testing Results** *(risk model performance metrics)*:
c-statistic not reported

2e.4 **If outcome or resource use measure is not risk adjusted, provide rationale:** Not applicable

2e
C☐ P☐ M☐ N☐ NA☐

2f. Identification of Meaningful Differences in Performance

2f.1 **Data/sample from Testing or Current Use** *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges

2f.2 **Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*:

2f
C☐ P☐ M☐ N☐

Comment [KP12]: 2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

Comment [k13]: 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP < 140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders) and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.

Comment [KP14]: 2d. Clinically necessary measure exclusions are identified and must be:
•supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;
AND
... [1]

Comment [k15]: 10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

Comment [KP16]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
•an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured out
... [2]

Comment [k17]: 13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and w
... [3]

Comment [KP18]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

Comment [k19]: 14 With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessatio
... [4]

Posterior probability distribution parameterized using the Gamma distribution

2f.3 Provide Measure Scores from Testing or Current Use *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:

5th	25th	Median	75th	95th
0.001361	0.002526	0.003658	0.005090	0.007724

2g. Comparability of Multiple Data Sources/Methods

2g.1 Data/sample *(description of data/sample and size)*: Not applicable

2g.2 Analytic Method *(type of analysis & rationale)*:
Not applicable

2g.3 Testing Results *(e.g., correlation statistics, comparison of rankings)*:
Not applicable

2g C☐ P☐ M☐ N☐ NA☐

Comment [KP20]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

2h. Disparities in Care

2h.1 If measure is stratified, provide stratified results *(scores by stratified categories/cohorts)*: Median income of patient´s ZIP code:
1) Estimate 2) Standard error 3) P-value: Relative to marked group-c 4) P-value:
2007 relative to 2006
First quartile (lowest income) 100.330 5.768 0.000 0.069
Second quartile 60.771 2.840 0.000 0.021
Third quartile 47.923 2.472 0.007 0.011
Fourth quartile (highest income)c 38.217 2.572 0.176

2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:
Users may stratify based on gender and race/ethnicity

2h C☐ P☐ M☐ N☐ NA☐

Comment [KP21]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);OR rationale/data justifies why stratification is not necessary or not feasible.

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*

2

Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, **met?**
Rationale:

2 C☐ P☐ M☐ N☐

3. USABILITY

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)

Eval Rating

3a. Meaningful, Understandable, and Useful Information

3a.1 Current Use: In use

3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*):
1) State of California:
http://www.oshpd.ca.gov/hid/products/preventable_hospitalizations/pdfs/PH_REPORT_WEB.pdf
2) State of New Jersey: Find and Compare Quality Care in New Jersey Hospitals,
http://www.nj.gov/health/healthcarequality/

3a C☐ P☐ M☐ N☐

Comment [KP22]: 3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives). An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.

3) Niagara Health Quality Coalition and Alliance for Quality Health Care: New York State Hospital Report Card, http://www.myhealthfinder.com/
4) State of Texas: Reports on Hospital Performance, http://www.dshs.state.tx.us/thcic/
5) Maine: Maine Health Data Organization: http://gateway.maine.gov/mhdo2008Monahrq/home.html
6) Hawaii: awaii Health Information Corporation: http://hhic.org/publicreports.asp
7) Nevada: Nevada Compare Care: http://www.nevadacomparecare.net/monahrq/home.html

In use as a part of the AHRQ Quality Indicators. They are reported in numerous forums including: http://hcupnet.ahrq.gov/HCUPnet.jsp?Id=EB57801381F71C41&Form=MAINSEL&JS=Y&Action=%3E%3ENext%3E%3E&_MAINSEL=AHRQ%20Quality%20Indicators

This measure is used in the Monahrq system that is provide for public reporting and quality improvement throughout the United States: http://monahrq.ahrq.gov/

3a.3 **If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):**
The software is publicly available free of charge (www.qualityindicators.ahrq.gov/). Users apply the software to their own administrative data (UB-04 or claims) that is readily available. Hundreds of users have downloaded AHRQ Quality Indicator software.

This measure is used in the Monahrq system that is provided for public reporting and quality improvement throughout the United States: http://monahrq.ahrq.gov/

Testing of Interpretability (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*)
3a.4 **Data/sample** *(description of data/sample and size)***:** AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges

3a.5 **Methods** *(e.g., focus group, survey, QI project)***:**
AHRQ has developed the Quality Indicators Mapping Tool to facilitate use of the Prevention Quality Indicators and incorporated the tool into the MONAHRQ software, which has undergone user beta testing and is now available for download.

3a.6 **Results** *(qualitative and/or quantitative results and conclusions)***:**
Several states including Maine, Hawaii and Nevada have begun public reporting using the MONAHRQ tool. See http://monahrq.ahrq.gov/

3b/3c. Relation to other NQF-endorsed measures

3b.1 **NQF # and Title of similar or related measures:**
None found.

(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**

3b. Harmonization — **3b**

If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population):
3b.2 **Are the measure specifications harmonized? If not, why?**
No competing measures found.

C☐ P☐ M☐ N☐ NA☐

3c. Distinctive or Additive Value — **3c**

3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:**
No competing measures found.

C☐ P☐ M☐ N☐ NA

5.1 **If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same**

Comment [KP23]: 3b. The measure specifications are harmonized with other measures, and are applicable to multiple levels and settings.

Comment [k24]: 16 Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., *influenza immunization* of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for *patients with diabetes*), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

Comment [KP25]: 3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare, is a more valid or efficient way to measure).

target population), Describe why it is a more valid or efficient way to measure quality: No competing measures found.	☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability*?	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, **met?** **Rationale:**	**3** **C**☐ **P**☐ **M**☐ **N**☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	**Eval Rating**
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	**4a** **C**☐ **P**☐ **M**☐ **N**☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) Yes 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.**	**4b** **C**☐ **P**☐ **M**☐ **N**☐
4c. Exclusions 4c.1 **Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** No 4c.2 **If yes, provide justification.**	**4c** **C**☐ **P**☐ **M**☐ **N**☐ **NA**☐
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences 4d.1 **Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** Coding professionals follow detail guidelines, are subject to training and credentialing requirements, peer review and audit. As a PQI, CHF is not a measure of hospital quality, but rather one measure of outpatient and other health care. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting. Some CHF care takes place in emergency rooms. As such, combining inpatient and emergency room data may give a more accurate picture of this indicator. Physician management of patients with congestive heart failure differs significantly by physician specialty. [1, 2] Such differences in community practices may be reflected in differences in CHF admission rates. [1] Edep ME, Shah NB, Tateo IM, et al. Differences between primary care physicians and cardiologists in management of congestive heart failure: relation to practice guidelines. J Am Coll Cardiol 1997;30(2):518-26. [2] Reis, SE, Holubkov R, Edmundowicz D, et al. Treatment of patients admitted to the hospital with congestive heart failure: specialty-related disparities in practice patterns	**4d** **C**☐ **P**☐ **M**☐ **N**☐
4e. Data Collection Strategy/Implementation	**4e**

Comment [KP26]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Comment [KP27]: 4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified and clinical data elements are specified for transition to the electronic health record.

Comment [KP28]: 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

Comment [KP29]: 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.

Comment [KP30]: 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:
This indicator is measured with high precision, and most of the observed variance reflects true differences across areas.
Risk adjustment for age and sex appears to affect the areas with the highest and lowest raw rates. Areas with high rates may wish to examine the clinical characteristics of their patients to check for a more complex case mix. Patient age, clinical measures such as heart function, and other management issues may affect admission rates.
As the causes for admissions may include poor quality care, lack of patient compliance, or problems accessing care, areas may wish to review CHF patient records to identify precipitating causes and potential targets for intervention.

4e.2 Costs to implement the measure (*costs of data collection, fees associated with proprietary measures*)**:**
All data necessary to calculate this measure are routinely collected for hospital administrative purposes. The software for calculating the measure is available for free at:
http://www.qualityindicators.ahrq.gov/software.htm

4e.3 Evidence for costs:
All data necessary to calculate this measure are routinely collected for hospital administrative purposes. The software for calculating the measure is available for free at:
http://www.qualityindicators.ahrq.gov/software.htm

4e.4 Business case documentation: All data necessary to calculate this measure are routinely collected for hospital administrative purposes. The software for calculating the measure is available for free at:
http://www.qualityindicators.ahrq.gov/software.htm

C☐ P☐ M☐ N☐

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility*? **4**

Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **4**
Rationale:
C☐ P☐ M☐ N☐

RECOMMENDATION

(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.** **Time-limited** ☐

Steering Committee: Do you recommend for endorsement?
Comments:
Y☐ N☐ A☐

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner)
Co.1 Organization
Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850

Co.2 Point of Contact
John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-

Measure Developer If different from Measure Steward
Co.3 Organization
Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850

Co.4 Point of Contact
John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-

Co.5 Submitter If different from Measure Steward POC John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-, Agency for Healthcare Research and Quality
Co.6 Additional organizations that sponsored/participated in measure development UC Davis, Stanford University, Battelle Memorial Institute
ADDITIONAL INFORMATION
Workgroup/Expert Panel involved in measure development **Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.** None
Ad.2 If adapted, provide name of original measure: None **Ad.3-5 If adapted, provide original specifications URL or attachment**
Measure Developer/Steward Updates and Ongoing Maintenance **Ad.6 Year the measure was first released:** 2001 **Ad.7 Month and Year of most recent revision:** 10, 2010 **Ad.8 What is your frequency for review/update of this measure?** Annual **Ad.9 When is the next scheduled review/update for this measure?** 05, 2011
Ad.10 Copyright statement/disclaimers: The AHRQ QI software is publicly available; no copyright disclaimers
Ad.11 -13 Additional Information web page URL or attachment:
Date of Submission (*MM/DD/YY*)**:** 02/01/2011

Page 14: [1] Comment [KP14] **Karen Pace** **10/5/2009 8:59:00 AM**

2d. Clinically necessary measure exclusions are identified and must be:

- supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;

AND

- a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus;

AND

- precisely defined and specified:
  - if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);

if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

Page 14: [2] Comment [KP16] **Karen Pace** **10/5/2009 8:59:00 AM**

2e. For outcome measures and other measures (e.g., resource use) when indicated:

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care;[Error! Bookmark not defined.] OR

rationale/data support no risk adjustment.

Page 14: [3] Comment [k17] **Karen Pace** **10/5/2009 8:59:00 AM**

13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

Page 14: [4] Comment [k19] **Karen Pace** **10/5/2009 8:59:00 AM**

14 With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74% v. 75%) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall poor performance may not demonstrate much variability across providers.