



## Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

### Brief Measure Information

**NQF #:** 3494

**Corresponding Measures:**

**Measure Title:** Hospital 90-Day, All-Cause, Risk-Standardized Mortality Rate (RSMR) Following Coronary Artery Bypass Graft (CABG) Surgery

**Measure Steward:** Centers for Medicare & Medicaid Services

**sp.02. Brief Description of Measure:** This measure estimates a hospital-level, risk-standardized mortality rate (RSMR) for patients 65 years and older discharged from the hospital following a qualifying isolated CABG procedure. Mortality is defined as death from any cause within 90 days of the procedure date of an index CABG admission. The measure was developed using Medicare Fee-for-Service (FFS) patients 65 years and older. An index admission is the hospitalization for a qualifying isolated CABG procedure considered for the mortality outcome. This measure may be used in one or more to be defined 90-day payment models.

**1b.01. Developer Rationale:** CABG is a priority area for outcome measure development because it is a common procedure associated with considerable morbidity, mortality, and health care spending. Between July 2013 to June 2016, there were 138,785 hospitalizations for CABG surgery among Medicare FFS patients in the U.S (Simoes et al., 2017). Acute myocardial infarction and coronary atherosclerosis represent the 5th and 9th most costly conditions in the U.S. across all ages and payers respectively; despite being relatively less common reasons for inpatient admission, these conditions are often treated by CABG procedures. The ranking for these conditions rises among patients 65 years and older (Torio et al., 2016). In fiscal year 2014, isolated CABG surgeries accounted for almost half (40.59%) of all cardiac surgery hospital admissions in Massachusetts (Massachusetts Data Analysis Center, 2014). In 2014, the average Medicare payment was \$32,499 for CABG without valve and \$45,873 for CABG plus valve surgeries (Drye et al., 2009; Pennsylvania Health Care Cost Containment Council, 2014). Mortality associated with CABG surgery occurs beyond 30 days and a longer period of follow-up naturally captures more events. Data from a Danish registry showed that the risk of mortality after CABG persists up to 90-120 days after surgery, when the risk plateaus (Siregar et al., 2013). All-cause observed 90-day mortality after isolated CABG is 4.8%, compared to an observed rate of 3.2% for the 30-day outcome period (Medicare Fee-For-Service data, July 1, 2013 – June 30, 2016). Furthermore, there was considerable variability in the observed overall 90-day mortality rate, with a 10th percentile of 6.9% and a 90th percentile of 18.8% mortality among hospitals. Variation in these rates suggests that there is room for improvement. Care following hospital discharge from the index operation impacts postoperative mortality, and this phenomenon may not be adequately captured within the 30-day timeframe. Postoperative readmission is common and is strongly associated with postoperative mortality (Greenblatt et al., 2012). In fact, 27% of Medicare beneficiaries who underwent surgical procedures experienced readmission within the 90-day period (Jencks et al., 2009). Several studies indicate that procedural factors (for example, wearable cardioverter defibrillators and use of extracorporeal circulation) and complications of care (for

example, post-surgical impairment of renal function) may be associated with mortality up to 90 days and beyond (Zishiri et al., 2013; Eifert et al., 2010; Brown et al., 2006). As the measurement period is extended, optimal postoperative management becomes increasingly important in reducing complications and averting deaths. For example, patients living farther away from the hospital, where they underwent their CABG surgery, were more likely to be readmitted to a different hospital and had an increased risk of mortality (Tsai et al., 2015). A 90-day CABG mortality measure that may be used in alternate payment models can reveal critical disparities such as this and help incentivize better post-operative care for all patients regardless of distance from a CABG surgery center. Hospitals are responsible for many post-acute care decisions and can further impact outcomes including mortality, by enhancing care coordination such as, admission to rehabilitation facilities and programs especially in the context of alternate payment models that incentivize coordinated post-discharge care.

#### References

- Brown JR, Cochran RP, Dacey LJ, et al. Perioperative increases in serum creatinine are predictive of increased 90-day mortality after coronary artery bypass graft surgery. *Circulation*. 2006;114(1 Suppl):I409-413.
- Drye E, Krumholz HM, Vellanky S, Wang Y. Probing new conditions and procedures for new measure development. 2009.
- Eifert S, Kilian E, Beiras-Fernandez A, Juchem G, Reichart B, Lamm P. Early and mid term mortality after coronary artery bypass grafting in women depends on the surgical protocol: Retrospective analysis of 3441 on- and off-pump coronary artery bypass grafting procedures. *Journal of Cardiothoracic Surgery*. 2010;5-90.
- Greenblatt DY, Greenberg CC, Kind AJ et al. Causes and implications of readmission after abdominal aortic aneurysm repair. *Ann Surgery*. 2012;256:595-605
- Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *The New England Journal of Medicine*. 2009;360:1418-28.
- Massachusetts Data Analysis Center. Adult coronary artery bypass graft surgery in the commonwealth of Massachusetts. <https://www.mass.gov/files/documents/2017/12/14/cabg-fy2014.pdf>. Accessed March 23, 2018.
- Pennsylvania Health Care Cost Containment Council. Hospital medicare payment. <http://www.phc4.org/reports/cabg/16/docs/Hospital%20Medicare%20Payment.pdf>. Accessed March 23, 2018.
- Simoes J, Grady J, DeBuhr J, et al. 2017 Procedure-specific measures updates and specifications report hospital-Level 30-Day risk-standardized mortality measures. <http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic/Page/QnetTier3&cid=1163010421830>. Accessed March 23, 2018.
- Siregar S, Groenwold RH, de Mol BA et al. Evaluation of cardiac surgery mortality rates: 30-day mortality or longer follow-up? *European Journal of Cardiothoracic Surgery*. 2013;44:875-83.
- Torio CM, Moore BJ. National inpatient hospital costs: The most expensive conditions by payer, 2013: Statistical Brief #204. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD) 2016.
- Tsai TC, Orav EJ, Jha AK. Care fragmentation in the postdischarge period: Surgical readmissions, distance of travel, and postoperative mortality. *JAMA Surgery*. 2015;150:59-64.
- Zishiri ET, Williams S, Cronin EM, et al. Early risk of mortality after coronary artery revascularization in patients with left ventricular dysfunction and potential role of the wearable cardioverter defibrillator. *Circ Arrhythm Electrophysiology*. 2013;6(1):117-128.

---

**sp.12. Numerator Statement:** The outcome for this measure is 90-day all-cause mortality. Mortality is defined as death for any reason within 90 days of the procedure date from the index admission for patients 65 and older discharged from the hospital after undergoing isolated CABG surgery.

**sp.14. Denominator Statement:** This claims-based measure can be used in the patient cohort aged 65 years or older.

The cohort includes admissions for patients who receive a qualifying isolated CABG procedure (see the attached Data Dictionary) and with a complete claims history for the 12 months prior to admission. CMS publicly reports this measure for those patients 65 years or older who are Medicare FFS beneficiaries admitted to non-federal hospitals.

If a patient has more than one qualifying isolated CABG admission in the measure period, the first CABG admission is selected for inclusion in the measure and the subsequent CABG admission(s) are excluded from the cohort.

**sp.16. Denominator Exclusions:** The 90-day CABG surgery mortality measure excludes index admissions for patients:

- 1) With inconsistent or unknown vital status or other unreliable data.
- 2) Who leave the hospital against medical advice (AMA).
- 3) With qualifying CABG procedures subsequent to another qualifying CABG procedure during the measurement period.

---

**Measure Type:** Outcome

**sp.28. Data Source:**

Claims

**sp.07. Level of Analysis:**

Facility

---

**IF Endorsement Maintenance – Original Endorsement Date:** 2019-10-24 12:33 PM

**Most Recent Endorsement Date:** 10/24/2019 12:33:53 PM

---

**IF this measure is included in a composite, NQF Composite#/title:**

**IF this measure is paired/grouped, NQF#/title:**

**sp.03. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?:**

## 1. Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria

**1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.**

**[Response Begins]**

No

**[Response Ends]**

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

**Current Submission:**

Updated evidence information here.

**Previous (Year) Submission:**

Evidence from the previous submission here.

**1a.01. Provide a logic model.**

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

**[Response Begins]**

**[Response Ends]**

**1a.02. Provide evidence that the target population values the measured outcome, process, or structure and finds it meaningful.**

*Describe how and from whom input was obtained.*

**[Response Begins]**

**[Response Ends]**

**1a.03. Provide empirical data demonstrating the relationship between the outcome (or PRO) and at least one healthcare structure, process, intervention, or service.**

**[Response Begins]**

**[Response Ends]**

**1b.01. Briefly explain the rationale for this measure.**

*Explain how the measure will improve the quality of care, and list the benefits or improvements in quality envisioned by use of this measure.*

**[Response Begins]**

CABG is a priority area for outcome measure development because it is a common procedure associated with considerable morbidity, mortality, and health care spending. Between July 2013 to June 2016, there were 138,785 hospitalizations for CABG surgery among Medicare FFS patients in the U.S (Simoes et al., 2017). Acute myocardial infarction and coronary atherosclerosis represent the 5th and 9th most costly conditions in the U.S. across all ages and payers respectively; despite being relatively less common reasons for inpatient admission, these conditions are often treated by CABG procedures. The ranking for these conditions rises among patients 65 years and older (Torio et al., 2016). In fiscal year 2014, isolated CABG surgeries accounted for almost half (40.59%) of all cardiac surgery hospital admissions in Massachusetts (Massachusetts Data Analysis Center, 2014). In 2014, the average Medicare payment was \$32,499 for CABG without valve and \$45,873 for CABG plus valve surgeries (Drye et al., 2009; Pennsylvania Health Care Cost Containment Council, 2014).

Mortality associated with CABG surgery occurs beyond 30 days and a longer period of follow-up naturally captures more events. Data from a Danish registry showed that the risk of mortality after CABG persists up to 90-120 days after surgery, when the risk plateaus (Siregar et al., 2013). All-cause observed 90-day mortality after isolated CABG is 4.8%, compared to an observed rate of 3.2% for the 30-day outcome period (Medicare Fee-For-Service data, July 1, 2013 – June 30, 2016). Furthermore, there was considerable variability in the observed overall 90-day mortality rate, with a 10th percentile of 6.9% and a 90th percentile of 18.8% mortality among hospitals. Variation in these rates suggests that there is room for improvement. Care following hospital discharge from the index operation impacts postoperative mortality, and this phenomenon may not be adequately captured within the 30-day timeframe. Postoperative readmission is common and is strongly associated with postoperative mortality (Greenblatt et al., 2012). In fact, 27% of Medicare beneficiaries who underwent surgical procedures experienced readmission within the 90-day period (Jencks et al., 2009). Several studies indicate that procedural factors (for example, wearable cardioverter defibrillators and use of extracorporeal circulation) and complications of care (for example, post-surgical impairment of renal function) may be associated with mortality up to 90 days and beyond (Zishiri et al., 2013; Eifert et al., 2010; Brown et al., 2006). As the measurement period is extended, optimal postoperative management becomes increasingly important in reducing complications and averting deaths. For example, patients living farther away from the hospital, where they underwent their CABG surgery, were more likely to be readmitted to a different hospital and had an increased risk of mortality (Tsai et al., 2015). A 90-day CABG mortality measure that may be used in alternate payment models can reveal critical disparities such as this and help incentivize better post-operative care for all patients regardless of distance from a CABG surgery center. Hospitals are responsible for many post-acute care decisions and can further impact outcomes including mortality, by enhancing care coordination such as, admission to rehabilitation facilities and programs especially in the context of alternate payment models that incentivize coordinated post-discharge care.

**References**

- Brown JR, Cochran RP, Dacey LJ, et al. Perioperative increases in serum creatinine are predictive of increased 90-day mortality after coronary artery bypass graft surgery. *Circulation*. 2006;114(1 Suppl):I409-413.
- Drye E, Krumholz HM, Vellanky S, Wang Y. Probing new conditions and procedures for new measure development. 2009.
- Eifert S, Kilian E, Beiras-Fernandez A, Juchem G, Reichart B, Lamm P. Early and mid term mortality after coronary artery bypass grafting in women depends on the surgical protocol: Retrospective analysis of 3441 on- and off-pump coronary artery bypass grafting procedures. *Journal of Cardiothoracic Surgery*. 2010;5-90.
- Greenblatt DY, Greenberg CC, Kind AJ et al. Causes and implications of readmission after abdominal aortic aneurysm repair. *Ann Surgery*. 2012;256:595-605
- Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *The New England Journal of Medicine*. 2009;360:1418-28.
- Massachusetts Data Analysis Center. Adult coronary artery bypass graft surgery in the commonwealth of Massachusetts. <https://www.mass.gov/files/documents/2017/12/14/cabg-fy2014.pdf>. Accessed March 23, 2018.
- Pennsylvania Health Care Cost Containment Council. Hospital medicare payment. <http://www.phc4.org/reports/cabg/16/docs/Hospital%20Medicare%20Payment.pdf>. Accessed March 23, 2018.

Simoes J, Grady J, DeBuhr J, et al. 2017 Procedure-specific measures updates and specifications report hospital-Level 30-Day risk-standardized mortality measures.  
<http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic/Page/QnetTier3&cid=1163010421830>. Accessed March 23, 2018.

Siregar S, Groenwold RH, de Mol BA et al. Evaluation of cardiac surgery mortality rates: 30-day mortality or longer follow-up? *European Journal of Cardiothoracic Surgery*.2013;44:875-83.

Torio CM, Moore BJ. National inpatient hospital costs: The most expensive conditions by payer, 2013: Statistical Brief #204. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD)2016.

Tsai TC, Orav EJ, Jha AK. Care fragmentation in the postdischarge period: Surgical readmissions, distance of travel, and postoperative mortality. *JAMA Surgery*. 2015;150:59-64.

Zishiri ET, Williams S, Cronin EM, et al. Early risk of mortality after coronary artery revascularization in patients with left ventricular dysfunction and potential role of the wearable cardioverter defibrillator. *Circ Arrhythm Electrophysiology*. 2013;6(1):117-128.

**[Response Ends]**

**1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.**

*Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.*

**[Response Begins]**

We conducted analyses using data from July 2014 to June 2017 Medicare claims data (n= 137,819 admissions from 1,183 hospitals).

In the hierarchical logistic regression model, each hospital has its own intercept (random intercept), which is used to measure the differences in mortality between hospitals while adjusting for case-mix (patient risk factors). The hospital-level risk standardized mortality rates (RSMRs) have a mean of 4.86% and range from 2.04-11.26% in the study cohort. As shown in Table 1, the median risk-standardized rate is 4.67% (25th and 75th percentiles are 4.08% and 5.49%, respectively).

Table 1: Distribution of Hospital-Level 90-Day CABG Risk-Standardized Mortality Rates (RSMRs)

Description//90-Day Risk-Standardized Mortality Rates (%)

Number of Hospitals//1183

Mean (SD)//4.86 (1.17)

100th percentile (Maximum)//11.26

99th percentile //8.30

95th percentile //7.00

90th percentile //6.41

75th percentile //5.49

50th percentile (Median)//4.67

25th percentile //4.08

10th percentile //3.59

5th percentile //3.29

1th percentile //2.75

0th percentile (Minimum)//2.04

**[Response Ends]**

**1b.03. If no or limited performance data on the measure as specified is reported above, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement. Include citations.**

**[Response Begins]**

N/A

**[Response Ends]**

**1b.04. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.**

*Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.*

**[Response Begins]**

Distribution of 90-day CABG RSMRs by Proportion of Dual Eligible Patients:

Data Source: Medicare FFS claims and Medical Beneficiary Summary File (MBSF) data

Dates of Data: July 2014 through June 2017

Distribution of 90-day CABG RSMRs by Proportion of Patients with AHRQ SES Index Scores:

Data Source: Medicare FFS claims and The American Community Survey (2009-2013) data

Dates of Data: July 2014 through June 2017

Tables 2 and 3 below demonstrates the distribution of 90-day CABG RSMRs by proportion of Dual Eligible patients and AHRQ SES Index Score, respectively.

Table 2. Distribution of 90-Day CABG RSMRs by Proportion of Patients with Social Risk (Dual Eligible)

Variation in RSMR across hospitals (with at least 25 cases) by proportion of patients with social risk//

Description//Dual Eligible

Bottom/Top Quartile//Bottom Quartile//Top Quartile

Social Risk Proportion (%)// (0 - 2.86)//(>= 8.68)

# of hospitals//259//256

100% Max//9.3//11.3

90%//6.5//6.4

75% Q3//5.4//5.4

50% Median//4.6//4.7

25% Q1//4.0//3.9

10%//3.4//3.5

0% Min//2.2//2.3

Table 3. Distribution of 90-Day CABG RSMRs by Proportion of Patients with Social Risk (AHRQ SES Index)

Variation in RSMR across hospitals (with at least 25 cases) by proportion of patients with social risk//

Description//AHRQ SES Index

Bottom/Top Quartile//Bottom Quartile//Top Quartile

Social Risk Proportion (%)// (0 - 8.63)//(>= 26.67)

# of hospitals//256//253

100% Max//9.3//11.3

90%//5.8//6.9

75% Q3//5.0//6.0

50% Median//4.4//5.1

25% Q1//3.8//4.3

10%//3.3//3.7

0% Min//2.0//2.7

**[Response Ends]**

**1b.05. If no or limited data on disparities from the measure as specified is reported above, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in above.**

**[Response Begins]**

N/A

**[Response Ends]**



## 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. Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.

**spma.01. Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.**

[Response Begins]

No

[Response Ends]

**spma.02. Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.**

For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.

*For example, specifications may have been updated based on suggestions from a previous NQF CDP review.*

[Response Begins]

N/A

[Response Ends]

**sp.01. Provide the measure title.**

*Measure titles should be concise yet convey who and what is being measured (see [What Good Looks Like](#)).*

[Response Begins]

Hospital 90-Day, All-Cause, Risk-Standardized Mortality Rate (RSMR) Following Coronary Artery Bypass Graft (CABG) Surgery

[Response Ends]

**sp.02. Provide a brief description of the measure.**

*Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).*

[Response Begins]

This measure estimates a hospital-level, risk-standardized mortality rate (RSMR) for patients 65 years and older discharged from the hospital following a qualifying isolated CABG procedure. Mortality is defined as death from any cause within 90 days of the procedure date of an index CABG admission. The measure was developed using Medicare Fee-for-Service (FFS) patients 65 years and older. An index admission is the hospitalization for a qualifying isolated CABG procedure considered for the mortality outcome. This measure may be used in one or more to be defined 90-day payment models.

[Response Ends]

**sp.04. Check all the clinical condition/topic areas that apply to your measure, below.**

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- *Surgery: General*

**[Response Begins]**

**[Response Ends]**

**sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.**

**[Response Begins]**

**[Response Ends]**

**sp.06. Select one or more target population categories.**

*Select only those target populations which can be stratified in the reporting of the measure's result.*

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- *Populations at Risk: Populations at Risk*

**[Response Begins]**

**[Response Ends]**

**sp.07. Select the levels of analysis that apply to your measure.**

*Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.*

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- *Clinician: Clinician*
- *Population: Population*

**[Response Begins]**

Facility

**[Response Ends]**

**sp.08. Indicate the care settings that apply to your measure.**

*Check ONLY the settings for which the measure is SPECIFIED and TESTED.*

**[Response Begins]**

Inpatient/Hospital

**[Response Ends]**

**sp.09. Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials.**

*Do not enter a URL linking to a home page or to general information. If no URL is available, indicate "none available".*

**[Response Begins]**

N/A

**[Response Ends]**

**sp.12. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.**

*Attach an excel or csv file; if this poses an issue, [contact staff](#). Provide descriptors for any codes. Use one file with multiple worksheets, if needed.*

**[Response Begins]**

No data dictionary/code table – all information provided in the submission form

**[Response Ends]**

For the question below: state the outcome being measured. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.13. State the numerator.**

*Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome).*

*DO NOT include the rationale for the measure.*

**[Response Begins]**

The outcome for this measure is 90-day all-cause mortality. Mortality is defined as death for any reason within 90 days of the procedure date from the index admission for patients 65 and older discharged from the hospital after undergoing isolated CABG surgery.

**[Response Ends]**

For the question below: describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.14. Provide details needed to calculate the numerator.**

*All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets.*

*Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.*

**[Response Begins]**

This is an all-cause mortality measure, therefore any death within 90 days of the index procedure date from the index hospitalization is included in the measure outcome. We identify deaths for Medicare FFS patients 65 years or older using the Medicare Enrollment Database (EDB).

Numerator time window: 90 days from the procedure date of index CABG procedure.

This outcome measure does not have a traditional numerator and denominator like a core process measure (e.g., percentage of adult patients with diabetes aged 18-75 years receiving one or more hemoglobin A1c tests per year); thus, we are using this field to define the outcome and to which hospital the outcome is attributed when there are multiple hospitalizations within a single episode of care.

Outcome Attribution:

Attribution of the outcome in situations where a patient has multiple contiguous admissions, at least one of which involves an index CABG procedure (i.e., the patient is either transferred into the hospital that performs the index CABG or is transferred out to another hospital following the index CABG) is as follows:

1) If a patient undergoes a CABG procedure in the first hospital and is then transferred to a second hospital where there is no CABG procedure, the mortality outcome is attributed to the first hospital performing the index CABG procedure and the 90-day window starts with the date of index CABG procedure.

Rationale: A transfer following CABG is most likely due to a complication of the index procedure and that care provided by the hospital performing the CABG procedure likely dominates mortality risk even among transferred patients.

2) If a patient is admitted to a first hospital but does not receive a CABG procedure there and is then transferred to a second hospital where a CABG is performed, the mortality outcome is attributed to the second hospital performing the index CABG procedure and the 90-day window starts with the date of index CABG procedure.

Rationale: Care provided by the hospital performing the CABG procedure likely dominates mortality risk.

3) If a patient undergoes a CABG procedure in the first hospital and is transferred to a second hospital where another CABG procedure is performed, the mortality outcome is attributed to the first hospital performing the index (first) CABG procedure and the 90-day window starts with the date of index CABG procedure.

Rationale: A transfer following CABG is most likely due to a complication of the index procedure, and care provided by the hospital performing the index CABG procedure likely dominates mortality risk even among transferred patients.

**[Response Ends]**

For the question below: state the target population for the outcome. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.15. State the denominator.**

*Brief, narrative description of the target population being measured.*

**[Response Begins]**

This claims-based measure can be used in the patient cohort aged 65 years or older.

The cohort includes admissions for patients who receive a qualifying isolated CABG procedure (see the attached

Data Dictionary) and with a complete claims history for the 12 months prior to admission. CMS publicly reports this measure for those patients 65 years or older who are Medicare FFS beneficiaries admitted to non-federal hospitals.

If a patient has more than one qualifying isolated CABG admission in the measure period, the first CABG admission is selected for inclusion in the measure and the subsequent CABG admission(s) are excluded from the cohort.

**[Response Ends]**

For the question below: describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.16. Provide details needed to calculate the denominator.**

*All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets.*

*Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.*

**[Response Begins]**

The measure includes index admissions for patients:

1. Having a qualifying isolated CABG surgery during the index admission;
2. Enrolled in Medicare FFS Part A and Part B for the 12 months prior to the date of the index admission, and enrolled in Part A during the index admission; and,
3. Aged 65 or over.

Isolated CABG surgeries are defined as those CABG procedures performed without the following concomitant valve or other major cardiac, vascular, or thoracic procedures:

- Valve procedures;
- Atrial and/or ventricular septal defects;
- Congenital anomalies;
- Other open cardiac procedures;
- Heart transplants;
- Aorta or other non-cardiac arterial bypass procedures;
- Head, neck, intracranial vascular procedures; or,
- Other chest and thoracic procedures

This cohort is defined using International Classification of Diseases, 9th Revision, Clinical Modification (ICD-09-CM) procedure codes and/or International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-Procedure Coding System [PCS]) procedure codes identified in Medicare Part A Inpatient claims data. To create a clinically coherent population for risk adjustment and in accordance with existing NQF-approved CABG measures and clinical expert opinion, the measure is intended to capture isolated CABG patients (i.e., patients undergoing CABG procedures without concomitant valve or other major cardiac or vascular procedures see exclusion). ICD-09-CM and ICD-10-PCS procedure codes that indicate a patient has undergone a non-isolated CABG procedure (CABG surgeries that occur concomitantly with procedures that elevate patients' mortality risk) and thus does not meet criteria for inclusion in the measure cohort are used to identify such patients for removal from the cohort.

The ICD-09-CM and ICD-10-PCS procedure codes are listed in the attached Data Dictionary.

**[Response Ends]**

**sp.17. Describe the denominator exclusions.**

*Brief narrative description of exclusions from the target population.*

**[Response Begins]**

The 90-day CABG surgery mortality measure excludes index admissions for patients:

- 1) With inconsistent or unknown vital status or other unreliable data.
- 2) Who leave the hospital against medical advice (AMA).
- 3) With qualifying CABG procedures subsequent to another qualifying CABG procedure during the measurement period.

**[Response Ends]**

**sp.18. Provide details needed to calculate the denominator exclusions.**

*All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.*

**[Response Begins]**

The CABG surgery mortality measure excludes index admissions for patients:

1. With inconsistent or unknown vital status or other unreliable demographic (age and gender) data.  
Rationale: We do not include stays for patients where the age (indicated in the claim) is greater than 115, where the gender (indicated in the claim) is neither male nor female, where the admission date (indicated in the claim) is after the date of death in the Medicare Enrollment Database, or where the date of death (in the Medicare Enrollment Database) occurs before the date of discharge but the patient was discharged alive (indicated in the claim).
2. Discharged against medical advice (AMA).  
Rationale: Providers did not have the opportunity to deliver full care and prepare the patient for discharge. This information is taken from the discharge disposition in the claim.
3. With more than one qualifying CABG surgery admission in the measurement period.  
Rationale: CABG procedures are expected to last for several years without the need for revision or repeat revascularization. A repeat CABG procedure during the measurement period likely represents a complication of the original CABG procedure and is a clinically more complex and a higher risk surgery. Therefore, we select the first CABG surgery admission for inclusion in the measure and exclude subsequent CABG surgery admissions (additional claims indicating a CABG procedure was performed within 30-days of the index CABG procedure) from the cohort.

**[Response Ends]**

**sp.19. Provide all information required to stratify the measure results, if necessary.**

*Include the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate. Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format in the Data Dictionary field.*

**[Response Begins]**

N/A

**[Response Ends]**

**sp.20. Is this measure adjusted for socioeconomic status (SES)?**

**[Response Begins]**

[Response Ends]

**sp.21. Select the risk adjustment type.**

*Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.*

[Response Begins]

Statistical risk model

[Response Ends]

**sp.22. Select the most relevant type of score.**

*Attachment: If available, please provide a sample report.*

[Response Begins]

Rate/proportion

[Response Ends]

**sp.23. Select the appropriate interpretation of the measure score.**

*Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score*

[Response Begins]

Better quality = Lower score

[Response Ends]

**sp.24. Diagram or describe the calculation of the measure score as an ordered sequence of steps.**

*Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.*

[Response Begins]

The measure estimates hospital-level, 90-day, all-cause, RSMRs for CABG surgery using a hierarchical logistic regression model. In brief, the approach simultaneously models data at the patient and hospital levels to account for variance in patient outcomes within and between hospitals (Normand and Shahian, 2007). At the patient level, it models the log-odds of mortality within 90 days of the procedure date using age, sex, selected clinical covariates, and a hospital-specific effect. At the hospital level, the approach models the hospital-specific effects as arising from a normal distribution. The hospital effect represents the underlying risk of mortality at the hospital, after accounting for patient risk. The hospital-specific effects are given a distribution to account for the clustering (non-independence) of patients within the same hospital (Normand and Shahian, 2007). If there were no differences among hospitals, then after adjusting for patient risk, the hospital effects should be identical across all hospitals. The RSMR is calculated as the ratio of the number of “predicted” deaths to the number of “expected” deaths at a given hospital, multiplied by the national observed mortality rate. For each hospital, the numerator of the ratio is the number of deaths within 90 days predicted based on the hospital’s performance with its observed case mix, and the denominator is the number of deaths expected based on the nation’s performance with that hospital’s case mix. This approach is analogous to a ratio of “observed” to “expected” used in other types of statistical analyses. It conceptually allows a specific hospital’s performance, given its case mix, to be compared to an average hospital’s performance with the same case mix. Thus, a lower rate indicates lower-than-expected mortality rates or better quality, while a higher rate indicates higher-than-expected mortality rates or worse quality. The “predicted” number of deaths (the numerator) is calculated by using the coefficients estimated by regressing the risk factors and the hospital-specific effect on the risk of mortality. The estimated hospital-specific effect is

added to the sum of the estimated regression coefficients multiplied by the patient characteristics. The results are log transformed and summed over all patients attributed to a hospital to get a predicted value. The “expected” number of deaths (the denominator) is obtained in the same manner, but a common effect using all hospitals in our sample is added in place of the hospital-specific effect. The results are log transformed and summed over all patients in the hospital to get an expected value. To assess hospital performance for each reporting period, we re-estimate the model coefficients using the years of data in that period.

This calculation transforms the ratio of predicted over expected into a rate that is compared to the national observed mortality rate. The hierarchical logistic regression models are described fully in the original 90-day CABG mortality measure methodology report (YNHHS/CORE, 2018).

References

Normand S-LT, Shahian DM. 2007. Statistical and clinical aspects of hospital outcomes profiling. *Statistical Science* 22(2): 206-226.

Yale New Haven Health System/Center for Outcomes Research & Evaluation (YNHHS/CORE). Hospital-Level 90-day All-Cause Risk-Standardized Mortality Rate (RSMR) Following Coronary Artery Bypass Graft (CABG) Surgery; Updated Measure Methodology Report. 2018.

**[Response Ends]**

**sp.27. If measure testing is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.**

*Examples of samples used for testing:*

- *Testing may be conducted on a sample of the accountable entities (e.g., hospital, physician). The analytic unit specified for the particular measure (e.g., physician, hospital, home health agency) determines the sampling strategy for scientific acceptability testing.*
- *The sample should represent the variety of entities whose performance will be measured. The [2010 Measure Testing Task Force](#) recognized that the samples used for reliability and validity testing often have limited generalizability because measured entities volunteer to participate. Ideally, however, all types of entities whose performance will be measured should be included in reliability and validity testing.*
- *The sample should include adequate numbers of units of measurement and adequate numbers of patients to answer the specific reliability or validity question with the chosen statistical method.*
- *When possible, units of measurement and patients within units should be randomly selected.*

**[Response Begins]**

N/A. This measure is not based on a sample or survey.

**[Response Ends]**

**sp.30. Select only the data sources for which the measure is specified.**

**[Response Begins]**

Claims

**[Response Ends]**

**sp.31. Identify the specific data source or data collection instrument.**

*For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.*



**[Response Begins]**

Data sources for the Medicare FFS measure:

Medicare Part A inpatient and Part B outpatient claims: This data source contains claims data for FFS inpatient and outpatient services including: Medicare inpatient hospital care, outpatient hospital services, as well as inpatient and outpatient physician claims for the 12 months prior to an index admission.

Medicare Enrollment Database (EDB): This database contains Medicare beneficiary demographic, benefit/coverage, and vital status information. This data source was used to obtain information on several inclusion/exclusion indicators such as Medicare status on admission as well as vital status. These data have previously been shown to accurately reflect patient vital status (Fleming et al., 1992).

The American Community Survey (2009-2013): We examined disparities in performance according to the proportion of patients in each hospital who were dual eligible for both Medicare and Medicaid insurances. We also used the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) index score derived from the American Community Survey (2009-2013) to study the association between our measure and SES.

Master Beneficiary Summary File (MBSF)

The MBSF is an annually created file that contains enrollment information for all Medicare beneficiaries, including dual eligible status. Years 2014-2017 were used.

The Society of Thoracic Surgeons (STS) CABG Composite Online Star Ratings

Empiric validity testing was performed using the publicly available measure score of the Society of Thoracic Surgery (STS) CABG Composite Online Star Rating, which combines several measures across quality domains to score hospitals from one (low quality) to three (high quality) stars (The Society of Thoracic Surgeons, 2017).

References

Fleming C., Fisher ES, Chang CH, Bubolz D, Malenda J. Studying outcomes and hospital utilization in the elderly: The advantages of a merged data base for Medicare and Veterans Affairs Hospitals. Medical Care. 1992; 30(5): 377-91.

The Society of Thoracic Surgeons. STS Public Reporting Online. CABG Overall Composite Score. 2017. Available at: [https://publicreporting.sts.org/search/cabg\\_report\\_card/hospital?title=&field\\_year\\_target\\_id=11&field\\_state\\_value=All](https://publicreporting.sts.org/search/cabg_report_card/hospital?title=&field_year_target_id=11&field_state_value=All). Accessed December 1, 2018.

**[Response Ends]**

**sp.32. Provide the data collection instrument.**

**[Response Begins]**

No data collection instrument provided

**[Response Ends]**

**2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).**

***Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:***

***Current Submission:***

*Updated testing information here.*

***Previous Submission:***

*Testing from the previous submission here.*

**[Response Begins]**

No

**[Response Ends]**

**2ma.02. Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).**

***Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:***

***Current Submission:***

*Updated testing information here.*

***Previous Submission:***

*Testing from the previous submission here.*

**[Response Begins]**

No

**[Response Ends]**

**2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?**

**[Response Begins]**

**[Response Ends]**

**2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.**

**Please update the Scientific Acceptability: Validity - Other Threats to Validity section.**

**Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.**

**[Response Begins]**

**[Response Ends]**

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.

- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the [Submitting Standards webpage](#).
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the [2021 Measure Evaluation Criteria and Guidance](#).

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

2a. Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;

AND

If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, 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).

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

### Definitions

Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements 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 of the measure score addresses precision of measurement (e.g., signal-to-noise).

Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

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

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Risk factors that influence outcomes should not be specified as exclusions.

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 percent v. 75 percent) 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 less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

#### Current Submission:

Updated testing information here.

#### Previous (Year) Submission:

Testing from the previous submission here.

#### 2a.01. Select only the data sources for which the measure is tested.

[Response Begins]

[Response Ends]

#### 2a.02. If an existing dataset was used, identify the specific dataset.

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

[Response Begins]

[Response Ends]

**2a.03. Provide the dates of the data used in testing.**

*Use the following format: "MM-DD-YYYY - MM-DD-YYYY"*

[Response Begins]

[Response Ends]

**2a.04. Select the levels of analysis for which the measure is tested.**

*Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.*

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- Clinician: Clinician
- Population: Population

[Response Begins]

[Response Ends]

**2a.05. List the measured entities included in the testing and analysis (by level of analysis and data source).**

*Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.*

[Response Begins]

[Response Ends]

**2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.**

*If there is a minimum case count used for testing, that minimum must be reflected in the specifications.*

[Response Begins]

[Response Ends]

**2a.07. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing.**

[Response Begins]

[Response Ends]

**2a.08. List the social risk factors that were available and analyzed.**

*For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.*

[Response Begins]

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.09 check patient or encounter-level data; in 2a.010 enter “see validity testing section of data elements”; and enter “N/A” for 2a.11 and 2a.12.

**2a.09. Select the level of reliability testing conducted.**

*Choose one or both levels.*

[Response Begins]

[Response Ends]

**2a.10. For each level of reliability testing checked above, describe the method of reliability testing and what it tests.**

*Describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?**

*For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, [NQF Measure Evaluation Criteria](#)).*

[Response Begins]

[Response Ends]

**2a.12. Interpret the results, in terms of how they demonstrate reliability.**

*(In other words, what do the results mean and what are the norms for the test conducted?)*

[Response Begins]

[Response Ends]

**2b.01. Select the level of validity testing that was conducted.**

[Response Begins]

[Response Ends]

**2b.02. For each level of testing checked above, describe the method of validity testing and what it tests.**

*Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2b.03. Provide the statistical results from validity testing.**

*Examples may include correlations or t-test results.*

[Response Begins]

[Response Ends]

**2b.04. Provide your interpretation of the results in terms of demonstrating validity. (i.e., what do the results mean and what are the norms for the test conducted?)**

[Response Begins]

[Response Ends]

**2b.05. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified.**

*Describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided in Importance to Measure and Report: Gap in Care/Disparities.*

[Response Begins]

[Response Ends]

**2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.**

*Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.*

[Response Begins]

[Response Ends]

**2b.07. Provide your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities.**

*In other words, what do the results mean in terms of statistical and meaningful differences?*

[Response Begins]

[Response Ends]

**2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.**

*Describe the steps—do not just name a method; what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.**

*For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).*

[Response Begins]

[Response Ends]

**2b.10. Provide your interpretation of the results, in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders), and how the specified handling of missing data minimizes bias.**

*In other words, what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis was conducted, justify the selected approach for missing data.*

[Response Begins]

[Response Ends]

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



**2b.11. Indicate whether there is more than one set of specifications for this measure.**

[Response Begins]

[Response Ends]

**2b.12. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications.**

*Describe the steps—do not just name a method. Indicate what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.**

*Examples may include correlation, and/or rank order.*

[Response Begins]

[Response Ends]

**2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.**

*In other words, what do the results mean and what are the norms for the test conducted.*

[Response Begins]

[Response Ends]

**2b.15. Indicate whether the measure uses exclusions.**

[Response Begins]

[Response Ends]

**2b.16. Describe the method of testing exclusions and what was tested.**

*Describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used?*

[Response Begins]

[Response Ends]

**2b.17. Provide the statistical results from testing exclusions.**

*Include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores.*

[Response Begins]

[Response Ends]

**2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.**

*In other words, the value outweighs the burden of increased data collection and analysis. Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion.*

[Response Begins]

[Response Ends]

**2b.19. Check all methods used to address risk factors.**

[Response Begins]

[Response Ends]

**2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.**

[Response Begins]

[Response Ends]

**2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.**

[Response Begins]

[Response Ends]

**2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.**

[Response Begins]

[Response Ends]

**2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.**

*Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of  $p < 0.10$  or other statistical tests; correlation of  $x$  or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).*

[Response Begins]

[Response Ends]

**2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.**

[Response Begins]

[Response Ends]

**2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.**

*Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.*

[Response Begins]

[Response Ends]

**2b.26. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter “N/A” for questions about the statistical risk model discrimination and calibration statistics.**

*Validation testing should be conducted in a data set that is separate from the one used to develop the model.*

[Response Begins]

[Response Ends]

**2b.27. Provide risk model discrimination statistics.**

*For example, provide c-statistics or R-squared values.*

[Response Begins]

[Response Ends]

**2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).**

[Response Begins]

[Response Ends]

**2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.**

*The preferred file format is .png, but most image formats are acceptable.*

[Response Begins]

[Response Ends]

**2b.30. Provide the results of the risk stratification analysis.**

[Response Begins]

[Response Ends]

**2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).**

*In other words, what do the results mean and what are the norms for the test conducted?*

[Response Begins]

[Response Ends]

**2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.**

*Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.*

[Response Begins]

[Response Ends]

### 3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

---

**3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.**

**[Response Begins]**

Coded by someone other than person obtaining original information (e.g., DRG, ICD-10 codes on claims)

**[Response Ends]**

**3.02. Detail to what extent the specified data elements are available electronically in defined fields.**

*In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.*

**[Response Begins]**

ALL data elements are in defined fields in electronic claims

**[Response Ends]**

**3.03. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using data elements not from electronic sources.**

**[Response Begins]**

**[Response Ends]**

**3.04. Describe any efforts to develop an eCQM.**

**[Response Begins]**

**[Response Ends]**

**3.06. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**

**[Response Begins]**

**[Response Ends]**

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

**3.07. Detail any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm),**

**Attach the fee schedule here, if applicable.**

**[Response Begins]**

N/A

**[Response Ends]**

## 4. Usability and Use

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

---

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.

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

### 4a.01. Check all current uses. For each current use checked, please provide:

Name of program and sponsor

URL

Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]

Payment Program

[Response Ends]

### 4a.02. Check all planned uses.

[Response Begins]

Payment Program

[Response Ends]

### 4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

*For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?*

[Response Begins]

This measure is not currently publicly reported or used in any accountability program because the measure is being submitted to the National Quality Forum (NQF) for initial endorsement.

[Response Ends]

### 4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

*A credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*

**[Response Begins]**

The measure may ultimately be used in one or more CMS programs, such as the Hospital Value-Based Purchasing Program.

**[Response Ends]**

**4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.**

*Detail how many and which types of measured entities and/or others were included. If only a sample of measured entities were included, describe the full population and how the sample was selected.*

**[Response Begins]**

N/A

**[Response Ends]**

**4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.**

**[Response Begins]**

N/A

**[Response Ends]**

**4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.**

**[Response Begins]**

N/A

**[Response Ends]**

**4a.08. Summarize the feedback obtained from those being measured.**

**[Response Begins]**

N/A

**[Response Ends]**

**4a.09. Summarize the feedback obtained from other users.**

**[Response Begins]**

N/A

**[Response Ends]**

**4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.**

**[Response Begins]**

N/A



**[Response Ends]**

**4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included). If no improvement was demonstrated, provide an explanation. If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**

**[Response Begins]**

This is a new measure and there is no information available on performance improvement. This measure is not currently used in a quality improvement program, but a primary goal of the measure is to provide hospitals with performance information necessary to implement focused quality improvement efforts. Once the measure is implemented, we plan to examine trends in improvements by comparing RSMR over time.

**[Response Ends]**

**4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.**

**[Response Begins]**

N/A

**[Response Ends]**

**4b.03. Explain any unexpected benefits realized from implementation of this measure.**

**[Response Begins]**

N/A

**[Response Ends]**

## 5. Comparison to Related or Competing Measures

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

---

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

### 5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

*(Can search and select measures.)*

**[Response Begins]**

**[Response Ends]**

### 5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

*(Can search and select measures.)*

**[Response Begins]**

**[Response Ends]**

### 5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

**[Response Begins]**

Conceptually related measures but different target population and focus:

NQF # 0230 - Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following acute myocardial infarction (AMI) hospitalization for patients 18 and older

Same target population but different measure focus/definition:

NQF # 0119 - Risk-Adjusted Operative Mortality for CABG (STS)

NQF # 2515 - Hospital 30-day, all-cause, unplanned, risk-standardized readmission rate (RSRR) following coronary artery bypass graft (CABG) surgery

NQF #2558 - Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following coronary artery bypass graft (CABG) surgery

Similar measure focus but different target population:

NQF # 0123- Risk-adjusted operative mortality for aortic valve replacement (AVR) + CABG surgery (STS)

NQF # 0122- Risk-adjusted operative mortality for mitral valve (MV) replacement + CABG surgery (STS)

NQF # 1502- Risk-adjusted operative mortality for MV repair + CABG surgery (STS)

**[Response Ends]**

### 5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

**[Response Begins]**

Yes

**[Response Ends]**

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

**[Response Begins]**

The target population is isolated CABG patients for the proposed 90-day CABG mortality measure and all of the above measures that have different measure focus but same target population. The clinical cohort exclusions are harmonized to the extent possible given the differences between clinical registry (STS) and administrative claims data. The exclusions are nearly identical to the STS measures' cohort exclusions with the exception of epicardial MAZE procedures; STS excludes these procedures from the registry-based CABG mortality measure cohort because the version of registry data used for measure development did not allow for differentiation of epicardial and open maze procedures. We did not include in our list of related measures any non-outcome (e.g., process) measures with the same target population as our measure. Our measure cohort was heavily vetted by clinical experts, a technical expert panel, and a public comment period. In addition, the related claims-based 30-day isolated CABG mortality and readmission measures, which utilize the same definition of isolated CABG as this 90-day mortality measure, were validated using clinical registry data (STS Cardiac Surgery Registry data for the readmission measure and New York State Cardiac Surgery Registry data for the mortality measure). Because this is an outcome measure, clinical coherence of the cohort takes precedence over alignment with related non-outcome measures. Furthermore, non-outcome measures are limited due to broader patient exclusions. This is because they typically only include a specific subset of patients who are eligible for that measure (for example, patients who receive a specific medication or undergo a specific procedure).

**[Response Ends]**

**5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.**

*Provide analyses when possible.*

**[Response Begins]**

This measure was specifically developed for and may be used in 90-day payment models. It is not intended to replace the 30-day CABG mortality measure in its current programmatic use or public reporting.

**[Response Ends]**

## Appendix

**Supplemental materials may be provided in an appendix.:**

No appendix

## Contact Information

**Measure Steward (Intellectual Property Owner):** Centers for Medicare & Medicaid Services

**Measure Steward Point of Contact:** Dollar-Maples, Helen, [helen.dollar-maples@cms.hhs.gov](mailto:helen.dollar-maples@cms.hhs.gov)

**Measure Developer if different from Measure Steward:** Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (CORE)

**Measure Developer Point(s) of Contact:** Dorsey, Karen, [karen.dorsey@yale.edu](mailto:karen.dorsey@yale.edu)



## Additional Information

**1. Provide any supplemental materials, if needed, as an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be collated one file with a table of contents or bookmarks. If material pertains to a specific criterion, that should be indicated.**

**[Response Begins]**

No appendix

**[Response Ends]**

**2. List the workgroup/panel members' names and organizations.**

*Describe the members' role in measure development.*

**[Response Begins]**

Technical Expert Panel Members:

Vinay Badhwar, MD, FACS, FACC, Chair, Public Reporting Task Force, The Society of Thoracic Surgeons; Professor & Chair, Department of Cardiovascular and Thoracic Surgery, West Virginia University

Araceli Carrera, DNP, RN, NP-C, Cardiothoracic Nurse Practitioner

Lee Fleisher, MD, Professor of Medicine; Chair of Anesthesiology and Critical Care

Renante Ignacio, MD, FACP, AGSF, CMD, Medical Director

Alexander Iribarne, MD, MS, Assistant Professor, Surgery, Geisel School of Medicine at Dartmouth; Assistant Professor, Health Policy and Clinical Practice, The Dartmouth Institute; Cardiac Surgeon; Director of Cardiac Surgical Research

Cristina Lisa, Patient

Jeffrey Jacobs, MD, Chair, Workforce on National Databases, The Society of Thoracic Surgeons; Professor, Surgery & Pediatrics, Johns Hopkins University; Deputy Director, Johns Hopkins All Children's Heart Institute

Michael Mack, MD, FACC, Cardiothoracic Surgeon; Medical Director of Cardiothoracic Surgery

Sean O'Brien, MS, PhD Statistical Director, The Society of Thoracic Surgeons; Associate Professor, Biostatistics, Duke University

Lawrence Sadwin, Patient

David Shahian, MD, Chair, Council of Quality, Research & Patient Safety, The Society of Thoracic Surgeons; Vice President of Massachusetts General Hospital Center for Quality and Safety; Professor, Surgery, Harvard University

Joyce Sinclair, Family Caregiver

Technical Work Group Members

Paul Kurlansky, MD, Associate Professor of Surgery, Columbia University

Arnar Geirsson, MD, Associate Professor of Surgery (Cardiac Surgery); Section Chief, Cardiac Surgery, Yale School of Medicine

**[Response Ends]**

**3. Indicate the year the measure was first released.**

**[Response Begins]**

**[Response Ends]**

**4. Indicate the month and year of the most recent revision.**

**[Response Begins]**

**[Response Ends]**

5. Indicate the frequency of review, or an update schedule, for this measure.

[Response Begins]

[Response Ends]

6. Indicate the next scheduled update or review of this measure.

[Response Begins]

[Response Ends]

7. Provide a copyright statement, if applicable. Otherwise, indicate "N/A".

[Response Begins]

N/A

[Response Ends]

8. State any disclaimers, if applicable. Otherwise, indicate "N/A".

[Response Begins]

N/A

[Response Ends]

9. Provide any additional information or comments, if applicable. Otherwise, indicate "N/A".

[Response Begins]

N/A

[Response Ends]