National Quality Forum
Neurology Endorsement Maintenance Project, Phase I

0437 STK 04: Thrombolytic Therapy

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of acute ischemic stroke patients who arrive at this hospital within 2 hours of time last known well for whom IV t-PA was initiated at this hospital within 3 hours of time last known well. This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-1: Venous Thromboembolism (VTE) Prophylaxis, STK-2: Discharged on Antithrombotic Therapy, STK-3: Anticoagulation Therapy for Atrial Fibrillation/Flutter, STK-5: Antithrombotic Therapy By End of Hospital Day 2, STK-6 Discharged on Statin Medication, STK-8: Stroke Education, and STK-10: Assessed for Rehabilitation) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Acute ischemic stroke patients for whom IV thrombolytic therapy was initiated at this hospital within 3 hours (less than or equal to 180 minutes) of time last known well.

Denominator Statement: Acute ischemic stroke patients whose time of arrival is within 2 hours (less than or equal to 120 minutes) of time last known well.

Exclusions: • Less than 18 years of age
• Length of Stay > 120 days
• Enrolled in clinical trials related to stroke
• Admitted for elective carotid intervention
• Time last known well to arrival in the emergency department greater than 2 hours
• Documented reason for not initiating IV thrombolytic

Adjustment/Stratification: No risk adjustment or risk stratification N/A Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: David Hackney; David Tirschwell; Daniel Labovitz; William Barsan; Gregory Kapinos; Gail Austin Cooney MD; Jolynn Suko   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-7; N-0

1a. Impact: H-7; M-0; L-0; I-0; **1b. Performance Gap:** H-5; M-2; L-0; I-0

Rationale: 1a. tremendous impact, driving much of acute stroke care forward 1b. clear evidence of underuse, but no real data reported on disparities for this measure; they don't collect this information. Alot of indirect evidence of disparities described

**why "Length of Stay > 120 days"? why "Time last known well to arrival in the emergency department greater than 2 hours"? Is it thought to be impossible to thrombolize within an hour in the ED? why "3h" as opposed to 4.5h and detailing the criteria for 3.0-4.5h window? I guess it is because the data for practice gap for this expnaded time-window is more parsimonious...

**1b. Performance gap consistent across multiple studies, including implementation of this measure (48.3% 4Q2009 to 71.7% 3Q2011

1c. Evidence (*based on decision logic*): Y-7; N-0   **IF** a **Health Outcome**, rationale supports: Y-0; N-2; NA-5

Quantity: H-5; M-2; L-0; I-0; **Quality:** H-6; M-1; L-0; I-0; **Consistency:** H-7; M-0; L-0; I-0

Rationale: not that many studies available studies are RCTs of high quality quite consistent for the 0-3 hour time window, which is what is assessed here

**Relatively low number of studies, some imbalance in key prognostic variables (Quality)

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-7; N-0

2a. Reliability: H-5; M-2; L-0; I-0; **2b. Validity:** H-7; M-0; L-0; I-0

Rationale: 2a. present data suggesting 98.1% overall agreement, which does sound good, but then presentation of data is confusing, with a numerator and denominator that are not really explained; a reabstraction method is used which sounds appropriate. 2b. validity is well described in the various subsections; there was a clear identification of a distribution of performance; no statistically testing results were presented (they describe "target analysis" but do not report any related results)

**2a1.8 Denominator exclusions: documented reason for not initiating therapy (RNIT) 2a2.3 Overall agreement 98.1% but Reason for Not Initiating Therapy is 76.7% 2b3.3 Frequency of exclusion for RNIT is low 0.95%

3. Usability: H-7; M-0; L-0; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale:

4. Feasibility: H-7; M-0; L-0; I-0

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

0437 STK 04: Thrombolytic Therapy
<u>Rationale</u>: Already in use, likely driving increased use nationwide **The timing aspect of the measure may make it dificult an dhard to capture
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-7; N-0</mark> <u>Rationale</u>: important intervention, backbone of drive to aggressively treat acute stroke though actual effectiveness may be overhyped, appropriate usage is likely an excellent marker of quality of care **This is superior to NQF#2022 which essentially is the same measure <u>Additional Comments/Questions</u>:
Workgroup Call Summary
Scientific acceptability • ***Clarify numerator/denominator definition:*** The developer explained that the denominator includes patients who arrive at the hospital within 2 hours of the time last known well, and the numerator includes patients who were given t-PA treatment within 3 hours of the time last known well. This gives the hospital at least 60 minutes to make a determination and begin treatment. • ***Is it really reasonable for all hospitals to be able to administer t-PA within 60 minutes of arrival?*** The developer noted that most of the patients who are treated are those who arrive in the first two hours (presumably they present sooner because their strokes are more severe). The developer also acknowledged that this measure does not capture all stroke patients, but argued that it captures the most sensitive stroke patients. • ***How does administration of t-PA in the 3.0-4.5 hour window impact performance on this measure?*** The developer clarified that this measure only examines t-PA administration within the first 3 hours of the time last known well—so patients given t-PA in the 3.0-4.5 hour window are not included in this measure. The developer also assured the workgroup members that while these patients are certainly important, other quality measures/tools can be used to measure performance for this subgroup. • ***Why are patients with LOS >120 days excluded?*** The developer explained that this is an artifact of all of their performance measures because of billing cycles for CMS and because their program requires measure submission on a quarterly basis. *Other points of discussion* • ***There was some discussion of the merits of this measure compared to measure #2022.*** Further explanation and discussion of competing measures (including this measure as compared to #2022) will be done in the in-person meeting.
Additional staff notes
• Reliability was tested at the data element level. • Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level. • Measure #2022 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0242 Stroke and Stroke Rehabilitation: Tissue Plasminogen Activator (t-PA) Considered

Status: Maintenance, Original Endorsement: May 01, 2007
Description: Percentage of patients aged 18 years and older with a diagnosis of ischemic stroke who arrive at the hospital within 4.5 hours of time last known well who were considered for t-PA administration
Numerator Statement: Patients who were considered for t-PA administration
Denominator Statement: All patients aged 18 years and older with a diagnosis of ischemic stroke who arrive at the hospital within 4.5 hours of time last known well
Exclusions: None
Adjustment/Stratification: No risk adjustment or risk stratification Not applicable We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.
Level of Analysis: Facility
Type of Measure: Process
Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry
Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:**
American Academy of Neurology
American College of Radiology
National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: William Barsan; David Hackney; David Tirschwell; Daniel Labovitz; Gail Austin Cooney MD; Gregory Kapinos; Jolynn Suko   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-7; N-0
1a. Impact: H-6; M-1; L-0; I-0; 1b. Performance Gap: H-7; M-0; L-0; I-0
Rationale: stroke in a major public health challenge report 79% noncompliance
**1b.2 Only 3-8% of potentially eligible patients receive TPA 1b.3 Used in 2007-2009 PQRI 78.2% of patients did not meet measure
**1a could have included more references from the usual AHA guidelines introduction paragraphs. Focus on high impact of tPA by demonstrating the potent effect of giving tPA to many AIS patients. 3rd and 4th paragraphs of 1a actually address 1b. Still unclear to me what "tPA considered" means and why measuring it this way (Is it because if you think about it, you are more likely to give it?) This is not a health outcome measure but a process measure, correct? 1b has a high quality evidence reference.
**1b2 - PQRI Results indicating 78% did not meet measure.
1c. Evidence (*based on decision logic*): Y-7; N-0   IF a Health Outcome, rationale supports: Y-1; N-1; NA-5
Quantity: H-2; M-5; L-0; I-0; Quality: H-1; M-6; L-0; I-0; Consistency: H-2; M-5; L-0; I-0
Rationale: same arguments for all the tPA measures, except that this measure includes patients presenting up to 4.5 hours out; only a couple of studies that include the 3-4.5 hr window.   The 0-3 hour time window has more studies, of high quality and report consistent benefit. the 3-4.5 hour window really only contains a few studies, those are of high quality, and generally consistent - except that recent IST-3 was just reported with less positive results for 3-4.5 hour time window (but slightly different patients), and the FDA has recently decided (though i've not seen it in print) NOT to approve tPA for the 3-4.5 hour time window
**Now that I read these references, I understand the rationale for this process measure and it does seem to be tangible to a link to better outcome.
**Qunatity: > 7 RCTs, 3 prospective studies Quality: Assuming quality as medium per the grade assigned to the body of evidence which indicates strong methods, consistent results and no heterogeneity Consistency - see rationale for quality

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-6; N-1
2a. Reliability: H-3; M-3; L-1; I-0; 2b. Validity: H-3; M-3; L-1; I-0
Rationale: 2a1.3--I worry that this would only be a documentation measure and would not really contribute to changes in treatment of patients with acute stroke. I also have concerns about asking for this information out to 4.5 hours when tPA is not approved for use in this time period. I know this is supposed to be paired with 2022 but I don't really see that it is.
**validity is drawn into question as relates to the 3-4.5 hr time window being included
**2b1. The measure does not address the exclusions in the ECASS3 trial which provided the bulk of data supporting use of tPA in the 3-4.5 hour window. tPA has not been approved by the FDA for use beyond 3 hours, making its use as a national quality measure problematic. The measure does not account for the time needed for satisfactory evaluation prior to tPA delivery. Thus a patient arriving one minute before the end of the treatment window counts as a failure just as a patient arriving hours prior to the end of the window.
**2a1.1 Numerator "Patients considered for TPA administration" 2a2.3 Reliability scores 100% 2b2.1 Expert panel assessed face validity 2b5 79% of patients did not meet measure
**2a1: for the denominator, how do we define all AIS patients? ICD-9 upon final diagnosis at the end of hospitalization (missing ED d/c patients), or based on sxs in the ED? 2a2 and 2b have convincing data

0242 Stroke and Stroke Rehabilitation: Tissue Plasminogen Activator (t-PA) Considered
**2a1&2a2: Per modelrate description on table 8 re: EHR measures specifications -
3. Usability: H-3; M-2; L-1; I-1   (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)   <u>Rationale:</u> As stated before, I worry that this will not lead to any useful changes in accountability or treatment.   **I might consider scoring this section even lower; the performance rate was very low, and it is not clear how that rates are improving or driving improved quality
4. Feasibility: H-3; M-3; L-1; I-0   (*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)   <u>Rationale:</u> 4.c The standard is based on duration of symptoms, which is inherently subject to error and inaccuracy. It depends on who noticed the symptoms and whether they noted the time, or the "duration" is a retrospective estimate.   **4a. Often the documentation is poor for the reason tPA was not given 4b. not clear to me all elements are in EHRs, though they have specified a plan 4c. unaware of any problems 4d. minimal response to question, but is is being used
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-3   <u>Rationale:</u> The measurement and reporting on "consideration of intervention" as opposed to "intervention" seems less direct and less useful. The inclusion of the non FDA approved use of tPA thru 4.5 hours is also problematic.   **As structured, treatment failures may meet current standard of care.   <u>Additional Comments/Questions</u>: Think that this needs to be harmonized with 0437; both are facility based measures

Workgroup Call Summary

Scientific acceptability
- ***Isn't the denominator (patients 18+ within 4.5 hours of symptom onset) problematic, given that the FDA has not approved t-PA administration after three hours post-stroke?*** One workgroup member suggested that t-PA up to 4.5 hours goes beyond what some hospitals are willing to consider.
- ***The structure of the measure does not account for the fact that patients may technically arrive within the 4.5-hour window, but possibly still not in time for providers to evaluate the use of t-PA.***

Usability
- ***Is this measure needed if there is already a measure looking at t-PA administration?*** The developer argued that this measure—since it goes out to the 4.5 hours post-stroke—does "add value" (note that the paired t-PA administered measure put forward by this developer [#2022] only goes out to the 3 hours post-stroke).

Feasibility
- ***Can these data (including identifying exclusions) really be pulled from electronic sources?*** One workgroup member suggested that it would not be too difficult to alter the notes template in EHRs and therefore allow this measure to be computed.
- ***Documentation can be resource intensive: it is really worthwhile to have to document when you know immediately that you will not treat?*** One member argued that it is not unreasonable to need write down why/why not treatment was given if an ischemic stroke patient comes in within a reasonable time frame post-stroke. But this member again noted that t-PA administered in the 3.0-4.5 hour window is not FDA-approved. Another member hypothesized that poor performance on this measure may reflect poor documentation rather than poor care.

Other points of discussion
- ***The developer clarified that this measure is actually specified at the facility level.***

2022 Stroke and Stroke Rehabilitation: Tissue Plasminogen Activator (t-PA) Initiated

Status: New Submission

Description: Percentage of all patients aged 18 years and older with a diagnosis of ischemic stroke who present within two hours of time last known well and who are eligible for t-PA, for whom t-PA was initiated within three hours of time last known well

Numerator Statement: Patients for whom t-PA was initiated within three hours of time last known well

Denominator Statement: All patients aged 18 years and older with a diagnosis of ischemic stroke who present within two hours of time last known well and who are eligible for t-PA

Exclusions: Documentation of medical reason(s) for not initiating Tissue Plasminogen Activator (t-PA) within three hours of time last known well (eg, contraindications, conditions that might lead to increased risk of bleeding or unfavorable outcomes, other medical reasons)
Contraindications*

- CT findings of intracranial hemorrhage, subarachnoid hemorrhage, or major infarct signs
- History of intracranial hemorrhage, brain aneurysm, vascular malformation, or brain tumor
- Internal bleeding (less than 22 days)
- IV or IA t-PA given at a transferring hospital
- No IV access
- Platelets less than 100,000, PTT greater than 40 sec after heparin use
- PT greater than 15 or INR greater than 1.7, or unknown bleeding diathesis
- Recent intracranial or spinal surgery, head trauma, or stroke (less than 3 months)
- Recent surgery/trauma (less than 15 days)
- Seizure with postictal residual neurological impairments
- Suspicion of subarachnoid hemorrhage
- Systolic blood pressure greater than 185 or diastolic blood pressure greater than 110 mm hg.
- Unable to determine eligibility

Warnings/Conditions that might lead to increased risk of bleeding or unfavorable outcomes*:

- Acute pericarditis
- Advanced age
- Diabetic hemorrhagic retinopathy or other ophthalmic bleeding
- Glucose less than 50 or greater than 400 mg/dl
- Hemostatic defects including those secondary to severe renal or hepatic disease
- Left heart thrombus
- Life expectancy less than 1 year or severe co-morbid illness
- Patient currently receiving oral anticoagulants (e.g. Warfarin sodium, Coumadin)
- Pregnancy
- Rapid improvement
- Septic thrombophlebitis or occluded AV cannula at seriously infected site
- Stroke severity – Too mild
- Stroke severity – Too severe (e.g., NIHSS greater than 22)
- Subacute bacterial endocarditis

*Lists harmonized with The Joint Commission measure.

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Facility

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement **Other organizations:** American Academy of Neurology
American College of Radiology
National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: David Hackney; David Tirschwell; Jolynn Suko; Daniel Labovitz; William Barsan; Gregory Kapinos; Gail Austin Cooney MD   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-6; N-1
1a. Impact: H-6; M-1; L-0; I-0; 1b. Performance Gap: H-6; M-1; L-0; I-0
Rationale: SZ as contraindication... List not perfectly compliant with guidelines from AHA and ACCP Why this 2h window?
1c. Evidence (*based on decision logic*): Y-6; N-1   IF a Health Outcome, rationale supports: Y-1; N-2; NA-4

2022 Stroke and Stroke Rehabilitation: Tissue Plasminogen Activator (t-PA) Initiated
Quantity: H-3; M-4; L-0; I-0; Quality: H-4; M-3; L-0; I-0; Consistency: H-3; M-3; L-0; I-1
<u>Rationale</u>: There is not evidence of consistency
**there are not very many studies, but they are high quality and consistently show benefit in the 0-3 hour time window
2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-7; N-0
2a. Reliability: H-6; M-1; L-0; I-0; 2b. Validity: H-4; M-2; L-1; I-0
<u>Rationale</u>: 2b1.8--the exclusions section includes "warnings" which are not valid for excluding patients, such as advanced age, stroke severity too severe.
**2a1. Large number of exclusions & exceptions, many subjective 2a2. Despite this reliability testing of sample almost 100% 2b2.2 Expert panel, face validity testing
3. Usability: H-4; M-2; L-0; I-1
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
<u>Rationale</u>: really did not have much to say on this, generic language used to answer questions
**3a1 PCPI not yet publicly reporting data 3b2 PCPI recommends use in QI
4. Feasibility: H-5; M-1; L-0; I-1
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
<u>Rationale</u>: I don't understand the claim to full EHR implementation otherwise nonspecific generic short replies
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-5; N-2
<u>Rationale</u>: other measures reviewed are better
**This is an important measure of physician performance in an under treated population
<u>Additional Comments/Questions</u>: Concern regarding number and subjectivity of exclusions and exceptions
Workgroup Call Summary
Scientific acceptability
• ***In the denominator exclusion section (2a1.8), some of the warnings/conditions listed may imply that patients with those conditions should not be treated.*** Developers noted that this is the same list that is used in measure #0437 (which is put forward by a different developer). [NOTE: The developer for #0437 informed workgroup members that this list is taken from the FDA labeling instructions.] The developer for this measure (#0242) noted that this list is a list of exceptions, not exclusions, explaining that the use of exceptions allows for clinical judgment; the developer also noted that this list is not meant to be inclusive or exhaustive, but is a list of examples of why a provider may choose not to administer t-PA. However, one workgroup member noted a concern that this list may still give the impression that patients with these conditions should not be given t-PA.
Usability
• ***The submission states that all data elements for this measure are available in EHRs—does this mean that no manual abstraction is necessary?*** The developer clarified that all of the data elements for this measure were available from EHRs for some of the sites in which they tested the measure. They acknowledged that this is not necessarily the case for all EHRs. ***NQF clarification***: *The main goal for item # 4b.1 as it is currently implemented is to get an idea of whether abstraction from paper records is necessary in order to obtain the data elements for a measure. However, different developers may answer the question differently; further, this question does not quite delve into the reality that data elements that are explicitly defined in one EHR may not be in a different EHR.*
Additional staff notes
• Reliability was tested at the data element level.
• Empirical evidence for validity was not provided; however, developers describe their systematic approach to assess face validity.
• Measure #0437 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

1952 Time to Intravenous Thrombolytic Therapy

Status: New Submission

Description: Acute ischemic stroke patients aged 18 years and older receiving intravenous tissue plasminogen activator (tPA) therapy during the hospital stay and having a time from hospital arrival to initiation of thrombolytic therapy administration (door-to-needle time) of 60 minutes or less.

Median time from hospital arrival to administration of intravenous tissue plasminogen activator (tPA) therapy in acute ischemic stroke patients aged 18 years and older.

Numerator Statement: Acute ischemic stroke patients aged 18 years and older receiving intravenous tissue plasminogen activator (tPA) therapy during the hospital stay and having a time from hospital arrival to initiation of thrombolytic therapy administration (door-to-needle time) of 60 minutes or less.

Median time from hospital arrival to administration of intravenous tissue plasminogen activator (tPA) therapy in acute ischemic stroke patients aged 18 years and older.

Denominator Statement: All acute ischemic stroke patients who received intravenous thrombolytic therapy within 4.5 hours of symptom onset.

Included populations: Discharges with an ICD-9-CM Principal Diagnosis Code for Acute Ischemic Stroke.
-Diagnosis for ischemic stroke ICD-9: 433.01, 433.10, 433.11, 433.21, 433.31, 433.81, 433.91, 434.00, 434.01, 434.11, 434.91, 436
-Diagnosis for ischemic stroke ICD-10:I6322, I6529, I63139, I63239, I63019, I63119, I63219, I6359, I6359, I6320, I6609, I6619, I6629, I6330, I6340, I6350, I678.

Exclusions: • Patients less than 18 years of age
• Patient stroke occurred while in hospital
• Patients received in transfer from the inpatient, or outpatient of another facility
• Patients who did not receive thrombolytic therapy within 60 minutes and had a reason for delay documented by a physician/advanced practice nurse/physician assistant as the cause for delay: social, religious, initial refusal, hypertension requiring aggressive control with intravenous medications, inability to confirm patients eligibility, or further diagnostic evaluation to confirm stroke for patients with hypoglycemia (blood glucose < 50); seizures, or major metabolic disorders, or management of concomitant emergent/acute conditions such as cardiopulmonary arrest, respiratory failure requiring intubation), or investigational or experimental protocol for thrombolysis.
• Clinical trial

Adjustment/Stratification: No risk adjustment or risk stratification Not Applicable. Not Applicable.

Level of Analysis: Facility, Population : National, Population : Regional, Population : State

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data : Registry

Measure Steward: American Heart Association/American Stroke Association **Other organizations:** Not applicable

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: David Hackney; David Tirschwell; Daniel Labovitz; Jolynn Suko; Gregory Kapinos; Gail Austin Cooney MD; William Barsan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-7; N-0
1a. Impact: H-7; M-0; L-0; I-0; 1b. Performance Gap: H-7; M-0; L-0; I-0
Rationale: very important; door to treatment times are too long, and if shortened can convincingly benefit patients
1c. Evidence (*based on decision logic*): Y-7; N-0   IF a **Health Outcome**, rationale supports: Y-1; N-2; NA-4
Quantity: H-7; M-0; L-0; I-0; Quality: H-7; M-0; L-0; I-0; Consistency: H-6; M-1; L-0; I-0
Rationale: consistency: "In the 1995 NINDS trial of patients treated between 0-3 hours after stroke symptom onset, shorter onset to treatment time was associated with increased odds of a good functional outcome.(1) In the ECASS III trial of patients treated between 3-4.5 hours after stroke symptom onset, shorter onset to treatment time was not associated with increased odds of a good functional outcome (p=0.21)"
**much good evidence to support premise that earlier treatment leads to better outcomes
**Evidence grade given on p.9

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-7; N-0
2a. Reliability: H-5; M-2; L-0; I-0; 2b. Validity: H-5; M-2; L-0; I-0
Rationale: 2a. kappa good, not great, at 0.72 2b. this measure seems to include the 3-4.5 hour time window, which is not FDA approved use, and the evidence supporting is lesser
**Shouldn't there be a risk stratification model for this measure, as it takes longer to initiate tPA in more complex patients? Centers seeing more complex patients may take longer to push tPA but they have to contemplate harder risk/benefit assessment in these complex scenarios with ambiguous contraindications... Also, centers with hyperacute MRI pathways will be dinged by this measure even if they remain compliant with the 0-3h window... Yes, the earlier the better, but sllightly more confident choices with DWI or a poised decision over a 10min discussion

1952 Time to Intravenous Thrombolytic Therapy

over the phone with the expert attending might not neccarily be worse (measured as lesser quality by this proposed measure) than a zealed rushed decision by the first responder stubbornly applying the guidelines and administering faster tPA in a non-thoroughly thought decision tailored to the individual patient... Just something to think about before we launch this war btw centers on how quickly we administer tPA...

**2a2.3 inter-rater reliability 0.72 2b2.3 Evaluated over 9 years/1600 hospitals in GWTG Stroke; 100% expert panel agreement

**2a1.8--Should there be other exclusions noted such as elevated INR, platelet count < 25,000, recent major surgery?

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

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale: seems this is a new measure, so we are not sure yet

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

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale: some of these timing data elements are not routinely collected

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-6; N-1

Rationale: Is this maintenance or a new measure? the inclusion of the 3-4.5 hour time window needs discussion I may be flexible on my final conclusion after discussion

Additional Comments/Questions: Time as the ultimate quality measure may not be just... As I said, judicious use of tPA may require an extra 5minutes to run the decision by the expert and these 5 minutes may lead to avoidance of a few complications that could maybe balance the toll in terms of lesser efficacy/efficiency.

Workgroup Call Summary

Importance

- ***Would a rush to treat lead to inappropriate treatment or errors because not all the facts have been gathered***? One member noted that there is no evidence from the literature that early treatment leads to more problems—and in fact, the evidence suggests that earlier treatment leads to better outcomes (several workgroup members agreed).

Scientific acceptability

- ***Are the exclusions to the measure appropriate?*** Members voiced approval on the common reasons for delay in treatment, although one member noted in the preliminary evaluation that other exclusions might be needed (e.g., elevated INR, platelet count < 25,000, recent major surgery).

- ***Would there be unintended consequences?*** One member noted that, while not likely, there is a chance that providers might make an excuse for delay and not treat at all, rather than get dinged for taking too much time to administer treatment.

- ***Does this measure accurately distinguish quality of care between hospitals?*** *(In other words, is a hospital that does better with this measure actually better than one that can administer t-PA a little less quickly?)* Other workgroup members agreed that this is a valid question, although they also supported the message behind the measure (i.e., that t-PA given earlier is better).

- ***The denominator includes stroke patients treated with t-PA up to 4.5 hours after stroke—but administration of t-PA between hours 3.0-4.5 is not FDA approved.*** The developers noted that this measure does not suggest that patients be treated in hours 3.9-4.5—but rather that, if the decision is made to treat with t-PA, the treatment should be done as soon as possible. ***NQF note****: NQF guidance regarding the need for drugs/devices in NQF-endorsed measures to be FDA approved for the target condition does not necessarily speak to the timing for use of a drug. If (as for t-PA administration) the FDA has approved the drug for one time frame but not another, the Committee should consider the evidence concerning the timing, if relevant for a particular measure.*

- ***Should this measure be risk-adjusted in some way?*** The developer noted that the measure was constructed in such a way as to allow exclusions for patients that may require more than the usual amount of time—and also noted that other types of risk-adjustment may inadvertently "adjust away" important differences in care delivery.

Additional staff notes

- Reliability was tested at the data element level.
- In addition to a face validity assessment, this measure was tested at both the data element level and the measure score level.

0438 STK 05: Antithrombotic Therapy By End of Hospital Day Two

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of ischemic stroke patients who had antithrombotic therapy administered by end of hospital day two (with the day of arrival being day 1). This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-1: Venous Thromboembolism (VTE) Prophylaxis, STK-2: Discharged on Antithrombotic Therapy, STK-3: Anticoagulation Therapy for Atrial Fibrillation/Flutter, STK-4: Thrombolytic Therapy, STK-6: Discharged on Statin Medication, STK-8: Stroke Education, and STK-10: Assessed for Rehabilitation) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Ischemic stroke patients who had antithrombotic therapy administered by end of hospital day two.

Denominator Statement: Ischemic stroke patients

Exclusions: •   Less than 18 years of age
-      Duration of Stay < 2 days
-      Length of Stay > 120 days
-      Comfort measures only documented on the day of or day after hospital arrival
-      Enrolled in clinical trials related to stroke
-      Admitted for elective carotid intervention
-      IV OR IA thrombolytic therapy administered at this hospital or within 24 hours prior to arrival
-      Documented reason for not administering antithrombotic therapy by end of hospital day 2

Adjustment/Stratification: No risk adjustment or risk stratification N/A Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association. Input was also provided by the U.S. Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the **Committee Reviewers:** David Hackney; David Tirschwell; Daniel Labovitz; Jolynn Suko; Gail Austin Cooney MD; Gregory Kapinos; William Barsan *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-5; N-2
1a. Impact: H-7; M-0; L-0; I-0; 1b. Performance Gap: H-3; M-2; L-2; I-0
Rationale: 1b. performance high, may be reaching a ceiling; also did not report (do not have) disparities data
**Not a huge gap cited - 3-4% depending on source
**1b2 Average 97%, low 95.9% 4Q2009 to high 97.8% 3Q2-11; similar findings in PCNASR (92% in 2005 to 96% in 2009) and GWTGSM (91.46% in 2003 to 97.4% in 2009)
**Again, why being so broad with this terminology "antithrombotic". They mean "aspirin 150-325mg". If patient receives DVT Px dose at day 1 or 2 with heparin SQ, it may be included as compliant with this measure whereas this regimen is inadequate to address the role hereby intended.
1c. Evidence (*based on decision logic*): Y-7; N-0   IF a **Health Outcome**, rationale supports: Y-1; N-2; NA-4
Quantity: H-6; M-1; L-0; I-0 Quality: H-7; M-0; L-0; I-0 Consistency: H-7; M-0; L-0; I-0
Rationale: only a couple of very large trials for acute stroke, but they were of high quality and had very consistent results
**Grade A Level 1 evidence

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-7; N-0
2a. Reliability: H-5; M-2; L-0; I-0; 2b. Validity: H-6; M-1; L-0; I-0
Rationale: 2a. reliability testing data not clearly presented 2b. no clear identification of meaningful differences; suggest these values overestimate performance due to stroke center self selection
**Curently re-tooling to fit QDM
**2a2.3 Reliability high, 97.2% 2b2.2 Initial assessment with focus group and survey; ongoing validity through analysis of feedback 2b2.3 72 submission/past year, primarily RNPATT; also about new anticoagulants

3. Usability: H-5; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: 3a1 Reported by JCO; also used for MU-EHR Incentive Program 3.2 Used by JCO for Primary Stroke Center Certification 3b.2 Low performance gap demonstrated

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

0438 STK 05: Antithrombotic Therapy By End of Hospital Day Two
(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented) Rationale:
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-6; N-1 Rationale: No significant performance gap demonstrated Additional Comments/Questions:
Workgroup Call Summary
Importance   ***Opportunity for improvement***: While there was general agreement among workgroup members that this measure is important, there was some concern about the high performance rate (98%). Workgroup members wondered how much improvement is still feasible and noted that in considering not continuing to recommend this measure for endorsement, the Committee must consider the possibility that gains might lost as well as the effort required for continued use of the measure. Developers noted that the performance gap statistics are derived from PQRS data and that reporting for this measure is voluntary. They argued that performance in non-reporting hospitals may not reach such a high level. Workgroup members, however, also reported a very high performance rate on this measure in the Get with the Guidelines data. ***NQF note***: One potential option for measures that are "topped out" is to recommend them for endorsement, but with <u>reserve status</u>. This designation would retain such measures in the NQF Portfolio (to be used periodically for monitoring), while also communicating to potential users that the measures no longer address high leverage areas for accountability purposes.
Additional staff notes
- Reliability was tested at the data element level.   - Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level.

0435 STK 02: Discharged on Antithrombotic Therapy

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of ischemic stroke patients prescribed antithrombotic therapy at hospital discharge. This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-1: Venous Thromboembolism (VTE) Prophylaxis, STK-3: Anticoagulation Therapy for Atrial Fibrillation/Flutter, STK-4: Thrombolytic Therapy,STK-5: Antithrombotic Therapy By End of Hospital Day 2, STK-6 Discharged on Statin Medication, STK-8: Stroke Education, and STK-10: Assessed for Rehabilitation) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Ischemic stroke patients prescribed antithrombotic therapy at hospital discharge

Denominator Statement: Ischemic stroke patients

Exclusions: •   Less than 18 years of age
•   Length of Stay > 120 days
•   Comfort measures only documented
•   Enrolled in clinical trials related to stroke
•   Admitted for elective carotid intervention
•   Discharged to another hospital
•   Left against medical advice
•   Expired
•   Discharged to home for hospice care
•   Discharged to a health care facility for hospice care
•   Documented reason for not prescribing antithrombotic therapy at discharge

Adjustment/Stratification: No risk adjustment or risk stratification N/A Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association. Input was also provided by the U.S. Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: David Hackney; David Tirschwell; Daniel Labovitz; Gail Austin Cooney MD; Gregory Kapinos; Jolynn Suko; William Barsan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-6; N-1
1a. Impact: H-7; M-0; L-0; I-0; 1b. Performance Gap: H-4; M-2; L-1; I-0
Rationale: very high performance already, how much more improvement is possible?
**1a3 Antithrombotic therapy should be prescribed at discharge following acute ischemic stroke to reduce stroke mortality and morbidity as long as no contraindications exist 1b.2 MC data 1998-2001 show only 74.2% d/c'd on ATT; JCO actual use of measure 2009-2011 showed low of 98.1% 4Q2009 to 99.2% 3Q2011; GWTG-Stroke finds aggregate rate 98% 1b.4 Disparities in care for ethnic minorities exist and are not assessed with this measure
**Significant improvement in measure over time - 2% performance gap in some reportin ghospitals.
1c. Evidence (*based on decision logic*): Y-7; N-0   IF a **Health Outcome**, rationale supports: Y-0; N-2; NA-5
Quantity: H-7; M-0; L-0; I-0; Quality: H-7; M-0; L-0; I-0; Consistency: H-7; M-0; L-0; I-0
Rationale: 1c.5 "Landmark trials primarily focused on aspirin" 1c.8 Aspirin is cost-effective; other anti-thrombotic drugs add to cost of care

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-7; N-0

2a. Reliability: H-7; M-0; L-0; I-0; 2b. Validity: H-5; M-2; L-0; I-0
Rationale: 2a. confusing presentation of reliability data 2b. does ceiling effect come into play?
**2a2.3 Overall agreement 97.61% (91-99.7%) 2b2.2 Original validity assessed via survey & focus group; ongoing validity determined through feedback from measure users 2b2.3 43 submissions past year, primarily Reason for Not Prescribing ATT at Discharge; also questions about new anti-thrombotic agents 2b3.3 10.15% of patients had RNPAD
**Slightly better definition for denominator compared to similar measure proposed by AMA-PCPI.

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

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: ceiling effect may reduce this criteria
**3a.2 Data available on JCO website 3.2 Data required to maintain JCO Primary Stroke Center designation - no information on what % of reporting institutions seek this designation - could be cause of low performance gap 3b.2 Only useful as QI if performance gap and none

0435 STK 02: Discharged on Antithrombotic Therapy

demonstrated

**Utilized in JCAHO stroke center certification, Paul Coverdell registry - in place since 2009 and many cycles of improvement evident.

---

4. Feasibility: <mark>H-6; M-1; L-0; I-0</mark>

(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)

<u>Rationale</u>: 4.b hospitals use EHR manual abstraction from paper records, or both. Path to EHR is not described.

**Already in use in many places

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-5; N-1</mark>

<u>Rationale</u>: though i worry about the ceiling effect reducing our power to detect meaningful differences...

**No performance gap demonstrated; compliance is very high

**Prefered over the similar measure from AMA-PCPI

<u>Additional Comments/Questions</u>: Should harmonize with PCPI measure 0325; same population and variables except 0325 includes TIAs

Could this be changed to measure use of ASA (not other ATT) in order to promote cost-effective care?

**Only note that I would have is that this is one measure where there has been significant improvement over the last 6-7 years with rates as high as 98%.

Workgroup Call Summary

Importance

- ***Opportunity for improvement***: Workgroup members again agreed that this measure reflects an important topic, but noted the high performance reported for the measure. They expressed concern that removal of endorsement would negatively impact patients. ***NQF note***: See summary under measure #0438.

Additional staff notes

- Reliability was tested at the data element level.
- Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level.
- Measure #0325 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0325 Stroke and Stroke Rehabilitation: Discharged on Antithrombotic Therapy

Status: Maintenance, Original Endorsement: May 01, 2007

Description: Percentage of patients aged 18 years and older with a diagnosis of ischemic stroke or transient ischemic attack (TIA) who were prescribed antithrombotic therapy at discharge

Numerator Statement: Patients who were prescribed antithrombotic therapy at discharge

Denominator Statement: All patients aged 18 years and older with a diagnosis of ischemic stroke or transient ischemic attack (TIA)

Exclusions: All patient that expired during inpatient stay are excluded.

Documentation of medical reason(s) for not prescribing antithrombotic therapy at discharge (eg, patients admitted for performance of elective carotid intervention, patient had stroke during hospital stay, other medical reason(s))

Documentation of patient reason(s) for not prescribing antithrombotic therapy at discharge (eg, patient is receiving comfort care only, patient left against medical advice, other patient reason(s))

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

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

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:**
American Academy of Neurology
American College of Radiology
National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: David Hackney; David Tirschwell; Daniel Labovitz; Gail Austin Cooney MD; Gregory Kapinos; Jolynn Suko; William Barsan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-5; N-2

1a. Impact: H-7; M-0; L-0; I-0; 1b. Performance Gap: H-5; M-2; L-0; I-0

Rationale: 1.b Difficult to interpret a "performance gap" because the validity of the recommendations is not well established (in my opinion). There are too many variations in the nature of presentation, and other medical conditions to do RCT's on each permutation. This means the quality of the evidence is lacking to be sure how many patients should be discharged on antithrombotic therapy. Thus, changing the portion of patients might not be a good thing.

**Did document room for improvement, but 53% non-compliance does not seem credible, and had implications for face validity later. There was also an unclear presentation of data, the 53% noncompliance on page 2 drops to about 15-17% on top of page 3, difference not commented on.

**1b.2 2007-2012 PCPI data shows 53.03% of patients did not meet measure (in contrast to high compliance in JCO 0435 that measures almost identical population)

**Same comments as for measure "tPA considered". They could have rearranged some of their evidence from 1a to 1b. "antithrombotic at discharge" should be replaced by "most judicious antithrombotic regimen for secondary prevention of ischemic stroke", because discharging on heparin subcut for DVT prophylaxis is not adequate for the purpose of preventing a second stroke. Just terminological correction.

1c. Evidence (*based on decision logic*): Y-5; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-1; NA-6

Quantity: H-2; M-4; L-1; I-0; Quality: H-2; M-4; L-1; I-0; Consistency: H-2; M-3; L-1; I-1

Rationale: There were relatively few RCT's and some of the evidence is derived from expert opinion, rather than data. Consistency was not addressed

**I know these are high, but they simply stated the ASA guideline (the only one they quote) does not give this information; seems somewhat non-repsonsive. Also, do not provide estimates of net benefit as requested, again somewhat non-responsive.

**Both quantity and quality of evidence is coming out of guideline recommendations indicating consensus and/or single RCT. Grading scale on p.5 indicates mostly grade B evidence.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-7; N-0

2a. Reliability: H-7; M-0; L-0; I-0; 2b. Validity: H-4; M-3; L-0; I-0

Rationale: Reliability high based on re-abstraction of charts and agreement reported.   Validity is likely appropriate, but partially questioned by the very low compliance rates (esp. compared to the near 100% rate in the similar JC measure). Also, compared to the JC measure, TIAs are included here, where the JC excluded due to lack of reliable identification of such cases via ICD-9 coding.

**2a1.9 Denominator exceptions seem subjective 2a2.3 Denominator exceptions found to be highly reliable 2b2.2 does not seem to be a process for ongoing assessment of validity 2b3.3 Exception rate 27%; reliability of exceptions 98.9% agreement

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

0325 Stroke and Stroke Rehabilitation: Discharged on Antithrombotic Therapy
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)   <u>Rationale</u>: a little hard to say based on the validity question
4. Feasibility: H-5; M-2; L-0; I-0
(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)   <u>Rationale</u>: I'm perhaps not familiar enough with claim of total presence in EHR to feel comfortable with the lack of detail provided   **Measure in use since 2009
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-7; N-0   <u>Rationale</u>: I would like to reserve final decision until discussion with group and a few process questions have been answered   <u>Additional Comments/Questions</u>: CMS data suggest poor performance on this measure with a 50% failure rate but GWTG data suggest >98% performance for all hospitals. Could it be that this measure no longer measures anything useful?   **I think this measure should be harmonized with JCO #435 because they measure almost identical populations and processes Need to understand why the performance gap in this measure is so different from that in JCO #435

Workgroup Call Summary

Importance

- ***Opportunity for improvement***: Workgroup members questioned the difference in performance rate compared to what is reported in measure #0435. The developer explained that the 53% performance rate included in their submission reflected 2008 PQRS data, and that the PQRS rate for 2010 was 83% (distributional statistics were not provided for 2010). They also noted that reporting of this measure in PQRS is voluntary and that about 24 percent of eligible professionals participated in 2010.

Scientific Acceptability

- ***One member questioned whether the relatively lower performance of this measure (compared to measure #0435) could be due to non-documentation in claims data.*** The developer was unable to confirm this hypothesis.
- ***Inclusion of TIA patients in the denominator***: Workgroup members noted that this measure includes TIA patients, but noted that the parallel Joint Commission measure (#0435) does not include TIA patients because ICD-9 codes are not reliable for identifying TIA.   The developer suggested that even though identification of TIA patients using ICD-9 codes may not be reliable, it is still important that such patients be included in the measure because the guidelines support antithrombotic therapy for TIA patients. Workgroup members emphasized the need to harmonize this measure with #0435.

Additional staff notes

- Reliability was tested at the data element level.
- Empirical evidence for validity was not provided; however, developers describe their systematic approach to assess face validity.
- Measure #0435 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0439 STK-06: Discharged on Statin Medication

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of ischemic stroke patients with LDL greater than or equal to 100 mg/dL, or LDL not measured, or who were on a lipid-lowering medication prior to hospital arrival who are prescribed statin medication at hospital discharge. This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-1: Venous Thromboembolism (VTE) Prophylaxis, STK-2: Discharged on Antithrombotic Therapy, STK-3: Anticoagulation Therapy for Atrial Fibrillation/Flutter, STK-4: Thrombolytic Therapy, STK-5: Antithrombotic Therapy By End of Hospital Day 2, STK-8: Stroke Education, and STK-10: Assessed for Rehabilitation) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Ischemic stroke patients prescribed statin medication at hospital discharge

Denominator Statement: Ischemic stroke patients with an LDL greater than or equal to 100 mg/dL, OR LDL not measured, OR who were on a lipid-lowering medication prior to hospital arrival.

Exclusions: •   Less than 18 years of age
•   Length of Stay > 120 days
•   Comfort measures only documented
•   Enrolled in clinical trials related to stroke
•   Admitted for elective carotid intervention
•   Discharged to another hospital
•   Left against medical advice
•   Expired
•   Discharged to home for hospice care
•   Discharged to a health care facility for hospice care
•   Documented reason for not prescribing statin medication at discharge

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable. Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association. Input was also provided by the U.S. Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the **Committee Reviewers:** David Hackney; David Tirschwell; Jolynn Suko; Daniel Labovitz; Gail Austin Cooney MD; Gregory Kapinos; William Barsan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-7; N-0
1a. Impact: H-6; M-1; L-0; I-0; 1b. Performance Gap: H-4; M-3; L-0; I-0
Rationale: 1.b Compliance with this recommendation is already high, so it cannot increase by much. there is evidence of disparities in treatment and subsequent stroke risk, but it is unclear whether discharge prescription rates are responsible.
**1b. no direct data on subgroups disparities, using the measure data, were reported; also, performance is improving and we may be approaching a ceiling in performance
**1b.2 Performance gap 89.5% in 4Q2009 to 94.9% in 3Q2011; similar results from other measurement sources
1c. Evidence (*based on decision logic*): Y-7; N-0   IF a **Health Outcome**, rationale supports: Y-2; N-1; NA-4
Quantity: H-6; M-1; L-0; I-0; Quality: H-2; M-5; L-0; I-0; Consistency: H-7; M-0; L-0; I-0
Rationale: unclear how much effect statins have on reducing the risk of recurrent stroke, or whether they are useful in patients with hemorrhagic stroke. Unclear when statins should be started
**Quality rated as moderate per 1.6c

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-7; N-0
2a. Reliability: H-6; M-1; L-0; I-0; 2b. Validity: H-6; M-1; L-0; I-0
Rationale: 2a. confusing presentation of reliability data, though suggest overall agreement is high
**2a1.8--SPARCL excluded patients with Atrial Fibrillation, Cardiac emboli and SAH. It would make sense for this measure to use the same exclusions.

3. Usability: H-7; M-0; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

4. Feasibility: H-6; M-1; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified*

0439 STK-06: Discharged on Statin Medication
4d. Data collection strategy can be implemented)   <u>Rationale</u>: Currently beeing re-tooled to meet QDM per JCAHO p.24
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-7; N-0</mark>   <u>Rationale</u>:   <u>Additional Comments/Questions</u>:
Workgroup Call Summary
Importance   •   ***Opportunity for improvement***: Workgroup members again agreed that this measure reflects an important topic, but noted the high performance rate reported in the measure submission.   **Scientific Acceptability**   •   ***Exclusions***: Workgroup members were confused about which patients are excluded from the measure. The developers confirmed that patients with atrial fibrillation are excluded and noted that providers must explicitly document why statins are not prescribed.
Additional staff notes
•   Reliability was tested at the data element level.   •   Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level.

2017 Stroke and Stroke Rehabilitation: Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) Reports

Status: New Submission

Description: Percentage of final reports for CT or MRI studies of the brain performed either:

In the hospital within 24 hours of arrival, OR

In an outpatient imaging center to confirm initial diagnosis of stroke, TIA or intracranial hemorrhage

For patients aged 18 years and older with either a diagnosis of ischemic stroke or transient ischemic attack (TIA) or intracranial hemorrhage OR at least one documented symptom consistent with ischemic stroke or TIA or intracranial hemorrhage that includes documentation of the presence or absence of each of the following: hemorrhage and mass lesion and acute infarction

Numerator Statement: Final reports of the initial CT or MRI that include documentation of the presence or absence of each of the following: hemorrhage and mass lesion and acute infarction

Denominator Statement: All final reports for CT or MRI studies of the brain performed either:

In the hospital within 24 hours of arrival, OR

In an outpatient imaging center to confirm initial diagnosis of stroke, TIA or intracranial hemorrhage.

For patients aged 18 years and older with either a diagnosis of ischemic stroke or TIA or intracranial hemorrhage OR at least one documented symptom consistent with ischemic stroke or TIA or intracranial hemorrhage

Exclusions: None

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Imaging/Diagnostic Study, Electronic Clinical Data : Registry

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement **Other organizations:** American Academy of Neurology

American College of Radiology

National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Risha Gidwani; Terry Richmond; Ramon Bautista *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-0; N-3

1a. Impact: H-1; M-1; L-1; I-0; **1b. Performance Gap:** H-1; M-0; L-2; I-0

Rationale: 1a. Stroke affects large numbers of people, but the link between neuroimaging and being able to improve outcomes was not well explained. 1b. The data regarding gap in performance is likely largely due to lack of reporting, rather than lack of neuroimaging. Additionally, it is not clear that increasing the number of providers who conduct imaging between 24 hours on the inpatient side or on the outpatient side would improve outcomes for ischemic stroke, as this is a large window of time/wrong setting for use of tPa.

1b) performance gap shown and is based on voluntary physician quality reportng system so this may not be representative, but there is a clear potential for improvement.

1a3 does not state which specific goal is being addressed although it can be inferred that accurate diagnosis is the goal. 1b2 does not state which specific performance gap has been identified.

1c. Evidence (*based on decision logic*): Y-0; N-3   IF a **Health Outcome**, rationale supports: Y-0; N-2; NA-1

Quantity: H-0; M-1; L-0; I-2; **Quality:** H-0; M-1; L-1; I-1; **Consistency:** H-0; M-0; L-0; I-3

Rationale: The developers do not discuss studies, but rather AHA/ASA guidelines. The level of evidence used to support each guideline recommendation was not clear -- the level of evidence was not linked to a specific guideline, but was discussed in an abstract way. It is strongly suggested that developers use different categorizations for grading level of evidence that do not place randomized evidence in the same category of quality as nonrandomized studies. The guideline that was quoted recommended "Rapid neuroimagining" rather than neuroimaging within 24 hours, and the discrepancy between the time frame suggested by the guideline and the time frame suggested by the developers was not addressed. Additionally, for section 1c.13, if was unclear to what definition the "A, A, B" were referring, as the previous section 1c.12 provides two definitions each for A, B and C.

**Quality - I am torn between rating this as moderate or insufficient - based on presentationof guidelines with evidence ranked as A,A,B but evidence is not provided. Grading for guidleines recommendation was noted to be Class 1 across the 3 criteria - indicating evidence and/or general consensus.

**AHA/ASA/ACR guidelines did not addressed quantity, quality and consistency of results across studies.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-1

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

2017 Stroke and Stroke Rehabilitation: Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) Reports
Rationale: Reliability was high in the test sample, but the number of sites (3) and number of charts per site (~30) resulted in a very low sample size.   With respect to validity, it would be beneficial to see a comparison of charts versus billing data -- given the specifics of documentation required to fulfill the numerator criteria, it is possible that billing codes do not have adequate sensitivity and specificity as compared to clinical documentation within the medical record.   Additionally, consensus among 18 experts regarding the face validity is necessary, but not sufficient, to establish the validity of this measure. **2b) face validity using focus panel of 26 experts (18 of whome votes on this measure). Shows varaibility and opportunity for perfomrance improvement **2a2.1 reliability testing only included 95 patients   2b. There is a serious issue with validity. CT head do not consistently detect acute cerebral infarcts. Also, a diagnosis of TIA does not depend on CT or MRI findings.
3. Usability: H-1; M-0; L-2; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*) Rationale: 3b- the information provided about use in QI programs is not accurate -- the Joint Commission does not use the time frame and specifications that are being put forth by these developers.   Additionally, I worry about the need to collect data from both the inpatient and outpatient setting to fulfill a singular measure -- this poses data collection and aggregation difficulties that were not discussed. **This measure can be easily misunderstood/misinterpreted by the public.
4. Feasibility: H-0; M-1; L-1; I-1
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*) Rationale: 4a. Developers do not discuss collection of data elements in locations that do not have electronic medical records. 4c. Lack of congruence between clinical documentation and billing codes was not discussed. 4d. No information was given to support the statement, "This measure was found to be reliable and feasible for implementation." **4c. There is a susceptibility to unintended consequences of this measure.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-0; N-3
Rationale: Need discussion/information on Quality and Consistency of this process measure - to convert to a yes. **The measure is misleading and prone to be misunderstood by the lay public. It has to be modified to account for the limitations of MRI and CT scans in diagnosing acute strokes and TIAs. Additional Comments/Questions:
Workgroup Call Summary

Importance (evidence):
- *The developers base this measure on a guideline, but there is no discussion given about the quantity, quality, or consistency of the evidence underlying this measure.*
- *What is the relationship between this measure and a desired health outcome?* One member noted that the literature supporting the link from documentation to improved processes to improved outcomes was not provided.
- *It is unclear what the three grades (A,A,B) in* section 1c.13 refer to.
- *Is there any evidence to show that not documenting (per this measure) has actually harmed patients?* The developer suggested that this is a valid inference but does not believe this has been documented.

Scientific acceptability
- *Workgroup members questioned the reliability testing in only 95 patients across three sites.* The developers noted that they had added additional information to their measure submission to demonstrate expanded reliability testing.
- *One workgroup member expressed a need for a comparison between the claims data and medical records to test/demonstrate the validity of the measure.*

Feasibility
- *Data collection:* Developers clarified that the data elements for this measure can be captured in claims data through the use of the specified CPT-II codes and that manual record abstraction is not required.

Other points of discussion
- *What is actually being measured?* Developers clarified that this measure examines whether or not the CT/MRI documentation actually records the presence or absence of three elements: hemorrhage, mass lesion, acute infarction, with the idea that this measure would support the use of a standardized format for reporting so that non-radiologists can know that a scan shows that treatment with a thrombolytic is safe. They clarified that this measure is NOT assessing whether the reading of the CT/MRI is correct.
- *One member noted that CT/MRI may not pick up certain conditions (e.g., lacunar infarcts/TIA).*

Additional staff notes

2017 Stroke and Stroke Rehabilitation: Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) Reports

- Reliability was tested at the data element level.
- Empirical evidence for validity was not provided; however, developers describe their systematic approach to assess face validity.

0434 STK-01: Venous Thromboembolism (VTE) Prophylaxis

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of ischemic or hemorrhagic stroke patients who received VTE prophylaxis or have documentation why no VTE prophylaxis was given on the day of or the day after hospital admission. This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-2: Discharged on Antithrombotic Therapy, STK-3: Anticoagulation Therapy for Atrial Fibrillation/Flutter, STK-4: Thrombolytic Therapy, STK-5: Antithrombotic Therapy By End of Hospital Day 2, STK-6 Discharged on Statin Medication, STK-8: Stroke Education, and STK-10: Assessed for Rehabilitation) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Ischemic or hemorrhagic stroke patients who received VTE prophylaxis or have documentation why no VTE prophylaxis was given on the day of or the day after hospital admission.

Denominator Statement: Ischemic or hemorrhagic stroke patients

Exclusions: •   Less than 18 years of age
•   Length of Stay < 2 days
•   Length of Stay > 120 days
•   Comfort measures only documented on day of or day after hospital arrival
•   Enrolled in clinical trials related to stroke
•   Admitted for elective carotid intervention

Adjustment/Stratification: No risk adjustment or risk stratification N/A Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association. Input was also provided by the U.S. Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Michael Kaplitt; Gwen Buhr; Jocelyn Bautista; Salina Waddy; Jolynn Suko; William Barsan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-2
1a. Impact: H-5; M-1; L-0; I-0; 1b. Performance Gap: H-4; M-1; L-0; I-1
Rationale: stroke is common and PE/DVT after stroke is not rare.
**While stroke is common, VTE is seen in only a small proportion of stroke patients (1%). Current performance is about 90%.
1c. Evidence (*based on decision logic*): Y-5; N-1   IF a **Health Outcome**, rationale supports: Y-1; N-1; NA-4
Quantity: H-4; M-2; L-0; I-0; Quality: H-3; M-3; L-0; I-0; Consistency: H-3; M-2; L-0; I-1
Rationale: for pharmacologic VTE prophylaxis the evidence is consistent, but for nonpharmacologic prophylaxis the evidence is not strong.
**Evidence regarding effectiveness of IPC devices in stroke patients is limited.
**Both quantity and quality of evidence is coming out of guideline recommendations indicating consensus and/or single RCT. Grading scale on p.5 indicates mostly grade B evidence.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-6; N-0
2a. Reliability: H-4; M-2; L-0; I-0; 2b. Validity: H-3; M-3; L-0; I-0
Rationale: There are a lot of questions in implementing the measure. Some of the exclusions are not consistent with the evidence.
**Measure relies primarily on chart abstraction from physician and nursing documentation.

3. Usability: H-3; M-3; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

4. Feasibility: H-3; M-3; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: The exclusions and exceptions may not be documented

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-6; N-0
Rationale:

Additional Comments/Questions: This measure overlaps with 0240. This one has more detailed and consistent evidence presented.

Workgroup Call Summary

Importance

0434 STK-01: Venous Thromboembolism (VTE) Prophylaxis

- *Impact*: While stoke is common, one member noted that VTE affects only 1% of stroke victims.
- *Opportunity for improvement*:   The performance rate for this measure is 92.4% as of Q32011. One workgroup member asked how to determine if this is high enough. *NQF clarification*: *NQF does not enforce an absolute threshold beyond which we consider that there is no more opportunity for improvement; instead, we ask Committee members to use their experience and expertise to decide if the possibility of even a small percentage increase in improvement is enough to make a measure worthwhile.*
- *Evidence for pharmacological or mechanical prophylaxis*: This measure allows for either pharmacological or mechanical prophylaxis. One member noted that while there is strong evidence for pharmacological prophylaxis, the evidence is less strong for mechanical—but this member also suggested that mechanical prophylaxis may be better than nothing.

Scientific Acceptability

- *Some exclusions are not evidence-based*. Developers note in their submission (section 2b3.2) that this measure is one of a set of stroke care measures and they want to keep the exclusions consistent across the set of measures.
- *Numerator includes patients given prophylaxis as well as those with documentation that no prophylaxis was given*. However, one workgroup member noted that this measure, as specified, would give the percentage of patients who were treated appropriately, regardless of whether they actually received the prophylaxis or whether they were appropriately deemed not to require prophylaxis.
- *Please explain choice A for VTE prophylaxis in the numerator (see section 2a1.3)*: This is the value used if recording that no prophylaxis was given (see above). The developer explained that if no prophylaxis was given, but there is no documented reason why, then the patient would be included in the denominator, but not in the numerator; if no prophylaxis was given, but there is a documented reason why, then the patient would be included in both the numerator and the denominator.
- *Is it valid to allow all reasons that may be put forward for not offering prophylaxis to exclude a patient from the measure?* One member argued that some reasons—even if documented—might be questionable. The developer asserted that there are very clear guidelines about what would be an acceptable reason for inclusion in the numerator.

Feasibility

- *Manual abstraction from paper records likely required*
- *Need for documentation*: One member noted that doctors don't always document why they are not using prophylaxis.

Additional staff notes

- Reliability was tested at the data element level.
- Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level.
- Measure #0240 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0240 Stroke and Stroke Rehabilitation: Deep Vein Thrombosis (DVT) Prophylaxis for Ischemic Stroke or Intracranial Hemorrhage

Status: Maintenance, Original Endorsement: May 01, 2007

Description: Percentage of patients aged 18 years and older with a diagnosis of ischemic stroke or intracranial hemorrhage who were administered DVT prophylaxis by the end of hospital day two

Numerator Statement: Patients who were administrated Deep Vein Thrombosis (DVT) prophylaxis by the end of hospital day two

Denominator Statement: All patients aged 18 years and older with a diagnosis of ischemic stroke or intracranial hemorrhage

Exclusions: All patients that expired during inpatient stay are excluded.

Documentation of medical reason(s) for not administering DVT Prophylaxis by end of hospital day 2 (eg, patient is ambulatory, patient already on warfarin or another anticoagulant, other medical reason(s))

Documentation of patient reason(s) for not administering DVT Prophylaxis by end of hospital day 2 (eg, patient is receiving comfort care only, patient left against medical advice, other patient reason(s))

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

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

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:**
American Academy of Neurology
American College of Radiology
National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Fred Tolin; Michael Kaplitt; Jocelyn Bautista; Gwen Buhr; Salina Waddy   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-1
1a. Impact: H-4; M-1; L-0; I-0; **1b. Performance Gap:** H-3; M-2; L-0; I-0
Rationale: 1a. While stroke is common, PE only affects about 1% of stroke patients.   1b. Current performance is 79-84% based on 2009-2010 PQRS data.
**Stroke is a high-impact condition and PE after stroke is a major complication. DVT prophylaxis is not universally prescribed.
**large number of patients and high cause of mortality high disparity in blacks
1c. Evidence (*based on decision logic*): Y-4; N-1   IF a **Health Outcome**, rationale supports: Y-1; N-0; NA-4
Quantity: H-4; M-1; L-0; I-0; **Quality:** H-2; M-3; L-0; I-0; **Consistency:** H-2; M-2; L-1; I-0
Rationale: Evidence for effectiveness of IPC devices in stroke patients is limited.
**Many RCTs but the events of PE or death or hemorrhage to make meaningful comparisons. The evidence is not convincing and the CPG language is 'consider' DVT prophylaxis.
**Nine studies were included in the systematic review within the Cochrane database.   ACCP Recommendation B: Methods strong, results inconsistent - RCTs, heterogeneity present

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-0
2a. Reliability: H-2; M-2; L-0; I-0; **2b. Validity:** H-1; M-2; L-1; I-0
Rationale: Measure relies heavily on manual chart abstraction from physician and nursing documentation.
**The exclusions are not specifically defined. It says 'other medical reasons' which is broad and can be defined differently by different clinicians.

3. Usability: H-2; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: Currently being used for quality improvement

4. Feasibility: H-2; M-2; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: Most of the information is available with high certainty. The thing that may not be available and depends on the thorough documentation by clinicians is the exclusions.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-3; N-1
Rationale: The evidence is not strong enough and the exclusions are not explicitly defined.
Additional Comments/Questions: The Measure Developer/Steward does not acknowledge competing NQF measure 434 in Section 5.1.

Workgroup Call Summary

0240 Stroke and Stroke Rehabilitation: Deep Vein Thrombosis (DVT) Prophylaxis for Ischemic Stroke or Intracranial Hemorrhage

Importance:

- ***Evidence for pharmacological or mechanical prophylaxis:*** As with measure #0434, one member noted that this measure allows for either pharmacological or mechanical prophylaxis, but noted that there is not strong evidence for mechanical prophylaxis.

Scientific acceptability

- ***Exclusions***: Although this measure is similar to measure #0434, this measure does not exclude patients admitted for elective surgical procedures (e.g., carotid intervention).

Additional staff notes

- Reliability was tested at the data element level.
- Empirical evidence for validity was not provided; however, developers describe their systematic approach to assess face validity.
- Measure #0434 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting. (Developers noted on the workgroup call that this measure is partially harmonized with #0434, but clarified that documentation underlying not administering prophylaxis is not captured in the measure and that those patients are not included in the numerator but are flagged as exceptions to the measure.)

0436 STK-03: Anticoagulation Therapy for Atrial Fibrillation/Flutter

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of ischemic stroke patients with atrial fibrillation/flutter who are prescribed anticoagulation therapy at hospital discharge. This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-1: Venous Thromboembolism (VTE) Prophylaxis, STK-2: Discharged on Antithrombotic Therapy, STK-4: Thrombolytic Therapy, STK-5: Antithrombotic Therapy By End of Hospital Day 2, STK-6 Discharged on Statin Medication, STK-8: Stroke Education, and STK-10: Assessed for Rehabilitation) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Ischemic stroke patients prescribed anticoagulation therapy at hospital discharge

Denominator Statement: Ischemic stroke patients with with documented atrial fibrillation/flutter.

Exclusions: •   Less than 18 years of age
•   Length of Stay > 120 days
•   Comfort measures only documented
•   Enrolled in clinical trials related to stroke
•   Admitted for elective carotid intervention
•   Discharged to another hospital
•   Left against medical advice
•   Expired
•   Discharged to home for hospice care
•   Discharged to a health care facility for hospice care
•   Documented reason for not prescribing anticoagulation therapy at discharge

Adjustment/Stratification: No risk adjustment or risk stratification N/A Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association. Input was also provided by the U.S. Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Salina Waddy; Gwen Buhr; Michael Kaplitt; Jocelyn Bautista   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-2
1a. Impact: H-4; M-0; L-0; I-0; 1b. Performance Gap: H-2; M-1; L-1; I-0

Rationale: Depends on definition of considerable. A roughly 5% gap, with roughly 95% average performance rate, seems fairly small. Since the reduction in risk is roughly 3.4% of patients, 3.4% of 5% that are not meeting performance would be roughly 0.17% reduction in strokes each year. But since 2.3 million people have Afib, that is roughly 4000 strokes that would be prevented each year, so that would make the number reasonably considerable.

**1a. Stroke is a leading cause of morbidity/disability, and about 20% of strokes result from atrial fibrillation. 1b. Average performance in 3rd quarter 2011 was 94%, leaving a small performance gap.

1c. Evidence (*based on decision logic*): Y-3; N-1   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-4
Quantity: H-4; M-0; L-0; I-0 Quality: H-3; M-1; L-0; I-0 Consistency: H-2; M-1; L-0; I-1

Rationale: There is some controversy over when to initiate anticoagulant therapy after acute stroke and over benefits of one anticoagulant over another.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-0
2a. Reliability: H-0; M-4; L-0; I-0; 2b. Validity: H-2; M-2; L-0; I-0

Rationale: exceptions are not spelled out. Exclusions are not consistent with the evidence.
**2a. Agreement rate with some data elements is only 83-85%. 2b. There is no information provided regarding the impact of exclusions.

3. Usability: H-3; M-1; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale: Average performance was already quite high at 93% when this measure was introduced in 2009, and has improved modestly to 95%.

4. Feasibility: H-2; M-2; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale: There is significant variability in stroke patients regarding the intensity of monitoring.

0436 STK-03: Anticoagulation Therapy for Atrial Fibrillation/Flutter
**exclusions may not be documented **Date elements regarding exceptions requires manual chart abstraction.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-0 Rationale: Additional Comments/Questions: This measure is very similar to 0241. They should both not be adopted. **This measure needs to be closely compared to competing measure 241.
Workgroup Call Summary

Importance
    • *Evidence*: One workgroup member noted that the medical evidence to support anticoagulation is not controversial. However, there was some discussion about the evidence around the timing of anticoagulant therapy.
    • *Opportunity for improvement*:   The performance rate for this measure is 95% as of Q32011. While quite high, workgroups members noted that the absolute numbers of people not being treated is still high and there is substantial under-treatment among minority populations.
 	
Scientific acceptability
    • *Under-diagnosis of atrial fibrillation*: One member noted that this measure would potentially miss many patients who should be treated.
    • *How is A-fib/flutter defined?* One member noted that some patients have very brief episodes of A-fib. The developer clarified that the measure includes any patient for whom A-fib is documented during the hospital stay or for whom there is any documentation of past history of A-fib or flutter.
    • *Which anticoagulants would meet the measure?* One member noted that the specifications are vague as to which anticoagulants can be used, noting that, in the past, Warfarin or Coumadin were the only agents available for oral use, but recently several new agents have been introduced, which have some advantages over Warfarin.
    • *Reliability testing*: One member noted that the percentage agreement results for the inter-rater reliability tests for certain data elements were relatively low (~82%, ~85%). The developer explained that the measure specifications for what can be included in the numerator may not be consistent with drugs that are newly-approved (e.g., use of Dabigatran was not captured in the numerator for approximately 9 months—this likely contributed to the lower agreement values between the raters). [Workgroup members noted that this may also impact the usability and feasibility of the measure.]
 	
Usability/Feasibility
    • *Drug availability*: Workgroup members noted that newer anticoagulants may not be equally available to all patients and that the monitoring needed when Coumadin is prescribed may be prohibitive for many patients, especially those in rural areas.

Additional staff notes

• Reliability was tested at the data element level.
    • Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level.
    • Measure #0241 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0241 Stroke and Stroke Rehabilitation: Anticoagulant Therapy Prescribed for Atrial Fibrillation at Discharge

Status: Maintenance, Original Endorsement: May 01, 2007

Description: Percentage of patients aged 18 years and older with a diagnosis of ischemic stroke or transient ischemic attack (TIA) with documented permanent, persistent, or paroxysmal atrial fibrillation who were prescribed an anticoagulant at discharge.

Numerator Statement: Patients who were prescribed an anticoagulant at discharge
Discharge refers to discharge from the acute care setting, whether patient received care in the emergency department or as an inpatient or a rehabilitation facility.

Denominator Statement: All patients aged 18 years and older with a diagnosis of ischemic stroke or transient ischemic attack (TIA) with documented permanent, persistent, or paroxysmal atrial fibrillation

Exclusions: All patients that expired during inpatient stay are excluded.
Documentation of medical reason(s) for not prescribing anticoagulant therapy at discharge (eg, other medical reason(s))
Documentation of patient reason(s) for not prescribing anticoagulant therapy at discharge (eg, patient is receiving comfort care only, patient left against medical advice, other patient reason(s))

Adjustment/Stratification: No risk adjustment or risk stratification No risk adjustment or stratitification We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

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

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:** American Academy of Neurology
American College of Radiology
National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Michael Kaplitt; Jocelyn Bautista; Gwen Buhr; Salina Waddy   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-2
1a. Impact: H-3; M-1; L-0; I-0; 1b. Performance Gap: H-3; M-1; L-0; I-0
Rationale: 1a. Stroke is a leading cause of morbidity and atrial fibrillation is a significant, treatable cause of stroke. 1b. 2010 performance was only 79%.
**Stroke is common and afib is a strong risk factor for stroke. Anticoagulant therapy reduces the risk and only half of people are receiving anticoagulant therapy.
1c. Evidence (*based on decision logic*): Y-2; N-2   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-3
Quantity: H-2; M-0; L-0; I-2; Quality: H-2; M-1; L-0; I-1; Consistency: H-2; M-1; L-0; I-1
Rationale: Insufficient information submitted to judge quantity of evidence. Evidence is not clear regarding how to treat intermediate risk patients.
**There are lots of rcts with consistent results for preventing future strokes in someone who has had a stroke.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-0
2a. Reliability: H-1; M-2; L-0; I-1; 2b. Validity: H-1; M-3; L-0; I-0
Rationale: Relies on manual chart abstraction to determine exceptions.
**I am concerned that the exceptions are not specifically defined and the exception rate was 46% for the project. Despite this there was high agreement on the exceptions.
**how was afib defined? # of beats?

3. Usability: H-2; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

4. Feasibility: H-0; M-4; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: limits exist for identifying disparity of treatment which would be important. EKG alone can markedly underdiagnose afib, not all hospitals are able to monitor with tele

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-0

0241 Stroke and Stroke Rehabilitation: Anticoagulant Therapy Prescribed for Atrial Fibrillation at Discharge

Rationale: there is a lot of evidence that addresses this issue

Additional Comments/Questions: I am concerned about the lack of firm definitions of the exceptions.

Workgroup Call Summary

Importance:

- *Evidence for the measure:* Workgroup members noted that this submission did not completely describe the evidence behind the measure. Developers agreed to add additional evidence to the measure submission.

Scientific acceptability:

- *Measure specifications*: Workgroup members requested the definition of atrial fibrillation. They also noted that atrial flutter was not included as part of the measure numerator and that TIA patients are included in the denominator. Workgroup members expressed particular concern with the inclusion of TIA patients in the measure. The developers noted that inclusion of TIA patients is supported by the guidelines; however, workgroup members were concerned with the validity of the TIA diagnosis.
- *Exceptions:* Workgroup members requested a list of the measure exceptions. They also questioned the exception rate of 46%. The developers clarified that the exception rate for the measure in the PQRS program was 15% and that the 46% exception rate applied to the testing project.

Usability & Feasibility

- *Manual abstraction*: One workgroup member expressed concern that manual chart abstraction would be necessary to identify exceptions.   *NQF note*: For claims data, the use of exceptions are documented via CPT-II codes.

Additional staff notes

- Reliability was tested at the data element level.
- Empirical evidence for validity was not provided; however, developers describe their systematic approach to assess face validity.
- Measure #0436 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0243 Stroke and Stroke Rehabilitation: Screening for Dysphagia

Status: Maintenance, Original Endorsement: May 01, 2007

Description: Percentage of patients aged 18 years and older with a diagnosis of ischemic stroke or intracranial hemorrhage who receive any food, fluids or medication by mouth (PO) for whom a dysphagia screening was performed prior to PO intake in accordance with a dysphagia screening tool approved by the institution in which the patient is receiving care

Numerator Statement: Patients for whom a dysphagia screening was performed prior to PO intake in accordance with a dysphagia screening tool approved by the institution in which the patient is receiving care

Denominator Statement: All patients aged 18 years and older with a diagnosis of ischemic stroke or intracranial hemorrhage who receive any food, fluids or medication by mouth (PO)

Exclusions: All patients that expired during inpatient stay are excluded

Exceptions:

Documentation of medical reason(s) for not performing a dysphagia screening prior to taking any foods, fluids or medication by mouth (eg, patient without any focal findings and not thought to be having a stroke when initially evaluated, other medical reason(s))

Documentation of patient reason(s) for performing a dysphagia screening prior to taking any foods, fluids or medication by mouth (eg, patient left against medical advice, other patient reason(s))

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

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

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:**
American Academy of Neurology
American College of Radiology
National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Jocelyn Bautista; Michael Kaplitt; Gwen Buhr; Salina Waddy   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-0

1a. Impact: H-3; M-1; L-0; I-0; **1b. Performance Gap:** H-3; M-1; L-0; I-0

Rationale: 1a. Dysphagia is an important cause of morbidity following stroke. 1b. Average performance in 2010 was 84%.

**Stroke is common and dysphagia after stroke is common. There is a performance gap according to the data submitted.

**Specifically regarding disparities, has there been an evaluation or has there been a study that identied a lack of a disparity...most likely the former.

1c. Evidence (*based on decision logic*): Y-4; N-0   IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-4

Quantity: H-1; M-3; L-0; I-0; **Quality:** H-1; M-3; L-0; I-0; **Consistency:** H-0; M-4; L-0; I-0

Rationale: Evidence is grade B.

**studies are nonrandomized, but there are a lot of them and they show consistent results.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-0

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

Rationale:

3. Usability: H-3; M-1; L-0; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale:

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

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale:

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-0

Rationale:

Additional Comments/Questions:

Workgroup Call Summary

Importance

- *Evidence:* Workgroup members noted that there is good evidence that swallowing difficulties can be under-

0243 Stroke and Stroke Rehabilitation: Screening for Dysphagia

recognized, particularly in the general community; however, they also noted that the different diagnostic approaches for dysphagia have different levels of sensitivity.

- *Type of stroke*: Workgroup members noted that there are different rates of dysphagia depending on the type of stroke, and that some types of stoke do not affect swallowing at all.
- ***Workgroup members noted the potential for high impact in terms of reducing the number of aspirations.***
- *Opportunity for improvement*: One member noted that the statistics on performance gap are based on PQRS data, which likely results in a selection bias that may inflate the numbers somewhat (i.e., those who report on this measure in PQRS may be more successful in meeting this measure than those who do not report on it).

Scientific Acceptability

- *Exclusions versus exceptions*: Workgroup members were confused about the difference between exclusions and exceptions. The developer clarified that they distinguish between the two, noting that exclusions are absolute and reflect patients who should never be included in the denominator, while exceptions are used to remove patients from the denominator when the patient does not receive service for appropriate reasons.   However, workgroup members raised concerns about whether use of exceptions allows physicians too much latitude. While the developer noted that past studies have shown that typically the exceptions used are valid, workgroup members noted that this information was not provided in the submission.

Additional staff notes

- Reliability was tested at the data element level.
- Empirical evidence for validity was not provided; however, developers describe their systematic approach to assess face validity.

0440 STK-08: Stroke Education

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of ischemic or hemorrhagic stroke patients with documentation that they or their caregivers were given stroke education materials. This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-1: Venous Thromboembolism (VTE) Prophylaxis, STK-2: Discharged on Antithrombotic Therapy, STK-3: Anticoagulation Therapy for Atrial Fibrillation/Flutter, STK-4: Thrombolytic Therapy,STK-5: Antithrombotic Therapy By End of Hospital Day 2, STK-6 Discharged on Statin Medication, and STK-10: Assessed for Rehabilitation) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Ischemic or hemorrhagic stroke patients with documentation that they or their caregivers were given educational material addressing all of the following:

1. Activation of emergency medical system
2. Need for follow-up after discharge
3. Medications prescribed at discharge
4. Risk factors for stroke
5. Warning signs and symptoms of stroke

Denominator Statement: Ischemic stroke or hemorrhagic stroke patients discharged home

Exclusions: •   Less than 18 years of age
•   Length of Stay > 120 days
•   Comfort measures only documented
•   Enrolled in clinical trials related to stroke
•   Admitted for elective carotid intervention

Adjustment/Stratification: No risk adjustment or risk stratification N/A Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association. Input was also provided by the U.S. Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Jocelyn Bautista; Gwen Buhr; Michael Kaplitt; Jack Scariano; Salina Waddy   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-1
1a. Impact: H-3; M-2; L-0; I-0; **1b. Performance Gap:** H-2; M-2; L-1; I-0
Rationale: Stroke is common, but the data submitted did not specifically speak to how stroke education has had a positive impact. The current performance leaves some room for improvement, but not a lot.
1c. Evidence (*based on decision logic*): Y-5; N-0   IF a **Health Outcome**, rationale supports: Y-2; N-0; NA-3
Quantity: H-4; M-1; L-0; I-0; **Quality:** H-0; M-5; L-0; I-0; **Consistency:** H-0; M-4; L-0; I-1
Rationale: Quality of evidence is moderate and report mixed results.
**Many studies were observational and small. The outcome measures are often not patient outcomes like recurrent stroke, rates of depression, or disability.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-1
2a. Reliability: H-2; M-2; L-1; I-0; **2b. Validity:** H-1; M-3; L-0; I-1
Rationale: There are exclusions not consistent with the evidence.

3. Usability: H-2; M-2; L-1; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale:

4. Feasibility: H-1; M-3; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale:

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-4; N-1
Rationale:
Additional Comments/Questions: It seems like the patients with risk factors for stroke, but not having yet had a stroke are the ones most in need of education.

0440 STK-08: Stroke Education
Workgroup Call Summary

Importance

- ***Opportunity for improvement***: Workgroup members noted a 10-30% gap, depending on the area of stroke education; however, they also noted that the information submitted did not include data on disparities in stroke education.
- ***Targeted population***: One workgroup member noted that this measure is targeted towards those who have already had a stroke, although some of the evidence presented reflects stroke education in the general population.
- ***Relationship to desired outcome***: Workgroup members voiced substantial concern about how stroke education is related to outcome and what the evidence suggests about stoke education and subsequent desired outcomes (e.g., reoccurrence or compliance with treatment, etc.).   They also agreed that providing education does not necessarily mean that the patient will understand it and be able to act on it. One member also noted some evidence suggesting that stroke education may be more efficacious if provided to the caregiver than if provided to the patient.

Additional staff notes

- Reliability was tested at the data element level.
- Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level.

1955 NIH Stroke Scale Recorded

Status: New Submission
Description: Percent of patients aged 18 and older with ischemic stroke, or stroke not otherwise specified, with an initial NIH Stroke Scale recorded.
Numerator Statement: Patients in whom a NIH Stroke scale test was measured, and a total score is recorded for these patients, as part of initial evaluation upon arrival at the hospital.
Denominator Statement: Patients with a final clinical diagnosis of ischemic stroke or stroke not otherwise specified.
Exclusions: • Patient is less than 18 years
• Stroke occurred while patient was an inpatient at the hospital
• Stroke symptoms resolved at time of presentation
• Patient underwent elective carotid intervention
Adjustment/Stratification: No risk adjustment or risk stratification Not Applicable. Not Applicable.
Level of Analysis: Facility, Population : National, Population : Regional, Population : State
Type of Measure: Process
Data Source: Administrative claims, Electronic Clinical Data : Registry
Measure Steward: American Heart Association/American Stroke Association **Other organizations:** Not Applicable.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Risha Gidwani; Terry Richmond; Ramon Bautista *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-3; N-0
1a. Impact: H-2; M-1; L-0; I-0; 1b. Performance Gap: H-3; M-0; L-0; I-0
Rationale: 1a) severity of stroke linked to mortality and cost   1b) variation by race, arrival mode, hospital charactersitics etc so potential for improvement
1c. Evidence (*based on decision logic*): Y-3; N-0   IF a **Health Outcome**, rationale supports: Y-0; N-1; NA-1
Quantity: H-0; M-3; L-0; I-0; Quality: H-2; M-0; L-1; I-0; Consistency: H-2; M-1; L-0; I-0
Rationale: Classes used to grade evidence had vague definitions -- it is strongly suggested to use of a well-accepted grading scale, such as GRADE or AHRQ scale. A Class I grade is noted as being "evidence and/or general agreement," which are quite different in terms of quality of evidence.   Some methodological comments: Developers state, "The correlation between the NIHSS at baseline and the measures at 90 days demonstrates predictive validity." In actuality, a high correlation between the NIHSS and 90-day mortality, or other scales at 90-day mortality would allow one to make this statement. Secondly, unweighted kappa statistics should be presented along with weighted kappa statistics (and note that the citation was incorrectly applied to this paragraph).
**important for treatment decisions, risk stratification, prediction. Benefits of use outweigh risks

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-1; N-1
2a. Reliability: H-1; M-2; L-0; I-0, 2b. Validity: H-0; M-1; L-0; I-1
Rationale: Developers note the measure has been "directly demonstrated to be useful in improving patient outcomes" but provide no explanation or citation for this statement.
**2a) total score accuracy with kappa 0.89. inter-rater reliability of individual items acceptable.   2b) Face validity- panel described. Mortality prediction c-stat 0.82 (very good) - Concern is that validity data are presented for patients 65 years and older but specificaton includes patients 18 years and older.
**2a2.3 does not take into account situations where non-neurologists have to perform the NIHSS. This may affect reliability.

3. Usability: H-0; M-1; L-1; I-1
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: 3a. Developers state "inclusion of this measure in a public reporting program has the potential to increase awareness of the importance of this measure in possible [sic] reducing 30 day mortality in acute ischemic stroke patients" -- an explanation of how this can/will occur is warranted. 3b. Is it not clear how recording the NIHSS can improve quality of care (GTWG was linked to QI, but how is the NIHSS component of GTWG linked to quality improvement?)
**This measure may be helpful for internal auditing and advancing stroke research but not for public usage.

4. Feasibility: H-0; M-2; L-1; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: NIHSS data can certainly be collected, but the ability to automatically aggregate and report these data was not discussed -- the lack of this explanation is reasonable, as NQF did not request this information, but it is an important component of performance measurement.
**4b) recognizes this as a limitation and trying to get all elements included in EHR - difficult to assess likelihood of this.
**Obtaining NIHSS though validated is not routinely performed in all stroke patients and will require training of healthcare workers, especially

1955 NIH Stroke Scale Recorded
non-neurologists.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-3; N-0</mark> **Rationale:** My conclusion is quite preliminary -- before making a final decision, I would like to know whether knowing the NIHSS score changes practice, and whether the NIHSS can be used for risk-adjustment. The latter is quite important, but may require a billing code to be used for large scale risk-adjustment purposes. **Additional Comments/Questions:**
Workgroup Call Summary

Importance (evidence):
    • ***What is the relationship between this measure and a desired health outcome?*** The developers—as well as members of the workgroup—noted that this scale is used to help inform the decision on whether or not to administer t-PA. The developers have added additional narrative to inform the relationship between scores on the NIH stroke scale and administration of t-PA.
 	
Scientific acceptability (reliability)
    • ***How valid/reliable is the scale when administered by non-neurologists***? Developers have added summaries of several studies that they argue demonstrate that the NIHSS can be performed by non-neurologists.
    • ***Data showing the relationship between the score on the scale and 30-day mortality was for patients 65+, but the measure is specified for patients 18+.*** Developers have added summaries for of the links between the NIHSS score and functional status/mortality in patients younger than 65.
 	
Usability
    • ***Should this measure be used in public reporting?*** It was clarified that it is not an individual's score on the scale that would be publicly reported, but rather a facility's aggregate percentage of patients 18+ with ischemic/NOS stroke who were evaluated with the NIHSS.
 	
Other points of discussion
    • The scale may miss a lot of brain stem strokes.

0244 Stroke and Stroke Rehabilitation: Rehabilitation Services Ordered

Status: Maintenance, Original Endorsement: May 01, 2007

Description: Percentage of patients aged 18 years and older with a diagnosis of ischemic stroke or intracranial hemorrhage for whom occupational, physical, or speech rehabilitation services were ordered at or prior to inpatient discharge OR documentation that no rehabilitation services are indicated at or prior to inpatient discharge

Numerator Statement: Patients for whom occupational, physical, or speech rehabilitation services were ordered at or prior to inpatient discharge OR documentation that no rehabilitation services are indicated at or prior to inpatient discharge

Denominator Statement: All patients aged 18 years and older with a diagnosis of ischemic stroke or intracranial hemorrhage

Exclusions: None

Adjustment/Stratification: No risk adjustment or risk stratification Not applicable We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team

Type of Measure: Process

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

Measure Steward: American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI) **Other organizations:**
American Academy of Neurology
American College of Radiology
National Committee for Quality Assurance

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Jordan Eisenstock; Jane Sullivan *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-0
1a. Impact: H-1; M-1; L-0; I-0; **1b. Performance Gap:** H-1; M-1; L-0; I-0

Rationale: 1a. large numbers, severity 1b. lower than expected by CPG's

**Impact Data --> Data presented support high incidence/impact of stroke and consequent disability; defines "stroke rehabilitation" and potential benefits of same. Data doesn't specifically address the impact of stroke rehabilitation. Performance Gap data --> -demonstrated discrepancy btw CPG recommendations and prevalence of survivors who rec'd OP stroke rehabilitation. -50.64% of patients were not offered rehab services in 2007-2011 (CMS PQRS) -approx 70% reporting rate (2007 & 2010) for eligible reporters -disparities reported regarding care received, degree of disability, functional status, & discharge destination by race & gender

1c. Evidence (*based on decision logic*): Y-2; N-0 IF a **Health Outcome**, rationale supports: Y-0; N-0; NA-2
Quantity: H-1; M-1; L-0; I-0; **Quality:** H-1; M-1; L-0; I-0; **Consistency:** H-1; M-1; L-0; I-0

Rationale: Linkages present Guidelines supportive Approximately 20 studies reviewed

**Evidence on efficacy of stroke rehab - CPG & systematic review (13 studies) & 6 additional studies that evaluate subacute stoke rehab. 19 RCTs of delayed transfer to rehab addressed multiple outcomes - conflicting results for LOS; differences btw pt. characteristics, intensity of therapy not well described; no studies assessing need for institutionalization reported differences w control AHA/ASA guideline - recommends use of comprehensive stroke units Meta analysis (Foley) strong evidence...associated with decreased mortality, morbidity, but not need for institutionalization CPG (Early Mgmt w Ischemic Stroke) AHA/ASA - class I - evidence for and general agreement that the treatment is useful & effective

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-0
2a. Reliability: H-2; M-0; L-0; I-0; **2b. Validity:** H-2; M-0; L-0; I-0

Rationale: 2a2. Good analytic methods 2b. Expert

**Numerator, denominator, target population, time window clear. Stratification recommendations clear. Validity expert panel - content validity & 30 day public comment period expert panel (n=18) - face validity, 5 point Likert Scale, 94.44 % agreement Performance Differences - 40% of eligible professionals reporting > 80% of eligible differences - less than optimal performance Reliability Data Sample - (AMA - PCPI) - 4 diverse practice sites with different medical records systems; 172 eligible charts btw 1-12/09; data analysis - % agreement, Kappa statistic couldn't be calculated --> numerator, denominator and overall - all > 95%

3. Usability: H-2; M-0; L-0; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale: Measure in use by several QI programs

**Measure used in CMS PQRS In use in QA programs; JCHO primary stroke certification program, AHA/ASA Get with Guidelines Program & CDC Paul Coverdell Registry

4. Feasibility: H-2; M-0; L-0; I-0

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified*

0244 Stroke and Stroke Rehabilitation: Rehabilitation Services Ordered
4d. Data collection strategy can be implemented)   <u>Rationale</u>: Data in medical records (EHR, etc.)
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-2; N-0</mark>   <u>Rationale</u>:   <u>Additional Comments/Questions</u>:
Workgroup Call Summary
  Workgroup members agreed on the importance of rehab services and no workgroup members voiced concerns about the measure. Note, however, that only 2 workgroup members had evaluated this prior to the workgroup call.
Additional staff notes
- Reliability was tested at the data element level.   - Empirical evidence for validity was not provided; however, developers describe their systematic approach to assess face validity.   - Numerator details section (2a1.3) notes that services can include any combination of physical, cognitive, behavioral, or speech services   - Measure #0441 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0441 STK-10: Assessed for Rehabilitation

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure captures the proportion of ischemic or hemorrhagic stroke patients assessed for or who received rehabilitation services during the hospital stay. This measure is a part of a set of eight nationally implemented measures that address stroke care (STK-1: Venous Thromboembolism (VTE) Prophylaxis, STK-2: Discharged on Antithrombotic Therapy, STK-3: Anticoagulation Therapy for Atrial Fibrillation/Flutter, STK-4: Thrombolytic Therapy, STK-5: Antithrombotic Therapy By End of Hospital Day 2, STK-6 Discharged on Statin Medication, and STK-8: Stroke Education) that are used in The Joint Commission's hospital accreditation and Disease-Specific Care certification programs.

Numerator Statement: Ischemic or hemorrhagic stroke patients assessed for or who received rehabilitation services.

Denominator Statement: Ischemic or hemorrhagic stroke patients.

Exclusions: •　　　Less than 18 years of age
- Length of Stay > 120 days
- Comfort measures only documented
- Enrolled in clinical trials related to stroke
- Admitted for elective carotid intervention
- Discharged to another hospital
- Left against medical advice
- Expired
- Discharged to home for hospice care
- Discharged to a health care facility for hospice care

Adjustment/Stratification: No risk adjustment or risk stratification N/A Not applicable, the measure is not stratified.

Level of Analysis: Facility, Population : National

Type of Measure: Process

Data Source: Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

Measure Steward: The Joint Commission **Other organizations:** The stroke measure set was developed in collaboration with the American Heart Association/American Stroke Association. Input was also provided by the U.S. Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry.

Workgroup Evaluation Results

The following evaluation ratings and comments are from the **Committee Reviewers:** Mary Van de Kamp; David Hackney; Tina Cronin; Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-0
1a. Impact: H-5; M-0; L-0; I-0 **1b. Performance Gap:** H-3; M-1; L-0; I-0
Rationale: 1.b.2 It is difficult to measure the performance gap because it is based on voluntary reporting and adherence to guidelines, with uncertainty as to how many eligible patients are represented or properly classified.

**As stated above, stroke is a leading cause of serious, long-term disability, associated with significant costs. The primary goal of rehabilitation is to prevent complications, minimize impairments, and maximize function. Evidence suggests that better clinical outcomes are achieved when post-acute stroke patients receive coordinated, multidisciplinary evaluation and intervention through the provision of rehabilitation services. Therefore, the need for rehabilitation services after an ischemic or hemorrhagic stroke should be assessed prior to discharge from the acute care setting. Healthcare organizations that track rehabilitation assessments for internal quality improvement purposes have seen an improvement in the measure rate over time. This measure is included in the FY 2015 CMS Hospital Inpatient Quality Reporting Program which will also promote improvements in quality at the national level. According to a 2011 report from the American Heart Association/American Stroke Association, racial disparities in stroke care exist and are more predominant among people < 65 years of age. Evidence of disparities in stroke care between minority groups and whites include: lack of knowledge about the risk factors for stroke; lack of awareness about stroke signs and symptoms and the need for urgent treatment; and, access to care respecting prevention services, acute stroke treatment, and rehabilitation. Differences in care are also related to the socioeconomic status of minorities, insurance coverage, cultural beliefs and attitudes, language barriers, immigration status, mistrust of the healthcare system, and the number of providers representing minority groups. These are all factors contributing to the quality of stroke care (Cruz-Flores, et al. 2011).

**1a. large numbers, severity 1b. multiple studies available

**Numerator, denominator, exclusions clear. Evidence for impact based on stroke prevalence, mortality & morbidity.   Evidence for benefits of multi-disciplinary stroke care. HC organizations that track this measure have seen improving rates over time. Rates from JCHO and PCDAS demonstrate high adherence but a slight performance gap (3-4%) Disparity in stroke care/outcomes by race, gender, & SES (some disparities are significant)

1c. Evidence (*based on decision logic*): Y-5; N-0   IF a **Health Outcome**, rationale supports: Y-1; N-0; NA-3

0441 STK-10: Assessed for Rehabilitation

Quantity: H-5; M-0; L-0; I-0 Quality: H-2; M-3; L-0; I-0 Consistency: H-3; M-2; L-0; I-0
Rationale: 41 studies noted. Overall, the body of evidence strongly supports that units providing stroke rehabilitation are associated with improved functional outcomes for patients. The impact of stroke rehabilitation on outcomes has been difficult to quantify due to problems with study design and methodology (e.g., lack of randomization, inappropriate control group selection, failure to blind assessors, difficulty in controlling for all possible confounders) detected by systematic review.   Furthermore, issues inherent to stroke rehabilitation, such as controlling for spontaneous neurological recovery, daily fluctuation in individual function, and difficulties measuring functional outcomes have challenged study designs. Pre-selection of patients and observer measurement bias are additional concerns when studying the impact of stroke rehabilitation on outcomes (Foley, et al, 2011).
**Multiple metaanalyses cited, included RCT's
**Quantity - EBRSR cites 2000 studies including 1078 RCTs & 5 metanalyses on the effectiveness of stroke rehab. + findings include: increased function, reduced mortality & institutional care. Quality/Consistency - inconsistencies in design, methodology make consensus difficult   No evidence of harm   Cost effectiveness is uncertain   EBRSR & AHA/ASA CPGs both strongly recommend

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-5; N-0
2a. Reliability: H-5; M-0; L-0; I-0; 2b. Validity: H-4; M-1; L-0; I-0
Rationale: specific data collection, use of a vendor for database maintenance and benchmarking, specific guidelines for inclusion and exclusion criteria are stated. This measure focuses on determining the proportion of ischemic or hemorrhagic stroke patients who were assessed for or who received rehabilitation services. Forty percent of stroke patients are left with moderate functional impairments and 15% to 30% with severe disability (Bates B, et al., 2005). These measure specifications are consistent with clinical practice guidelines that recommend for an initial assessment of complications, impairment and rehabilitation needs, in addition to obtaining the medical history and physical examination, and assessment of stroke severity.
**2a2. Re-abstraction performed and high agreement rate
**Numerator, denominator, exclusions, target population clear. Sampling methodology, data analysis clear. JCHO data for 2010-1 show high agreement rate (98.3%)   Measure has been in use since 2009 Validity initially assessed via survey & focus groups of hospitals; since implementation validity assessed via analysis of user feedback (72 submissions) and expert technical advisory panel   some concern about exclusions relative to attainment of 100%   Voluntary data collection may result in overstating of % (relative to real performance)

3. Usability: H-5; M-0; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: 3a/3b- The Joint Commission has a longstanding commitment to providing meaningful information about the comparative performance of accredited organizations to the public. The Quality Check® Web site, www.qualitycheck.org, launched in 1996, fulfills this commitment. Among other things, Quality Check allows consumers to view or download free hospital performance measure results.
        Measure rates for STK-10 are included in the hospital performance measure results reported on Quality Check®. The Paul Coverdell National Acute Stroke Registry (PCNASR) also collects and publicly reports STK-10 measure data. Established by Congress in 2001, PCNASR is funded by the Centers for Disease control and Prevention (CDC) through a cooperative agreement with state health departments. The state health departments work with participating hospitals to track the care of hospitalized stroke patients to improve the quality of acute stroke care from the onset of stroke through hospital discharge. This STK-10 Assessed for Rehabilitation measure is included among the 15 clinical quality measures required in Stage 1 Meaningful Use of the Electronic -Health Record (EHR) Incentive Program for eligible hospitals and CAHs. This measure also will become a component of CMS's Hospital Inpatient Quality Reporting (IQR) for FY 2015 with data collection to begin January 2013.
**3a. Part of stage 1 meaningful use 3b. Aggregate scores improving suggesting use by hospitals to address need
**Rates included in Quality Check Website, PCNASR, EHR incentive program for hospitals & CAHs   Will become part of CMS Inpatient Quality Reporting   Measure specifications standardized & regularly updated; JCHO has process for fb   Data suggests improvements over time

4. Feasibility: H-3; M-2; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: 4.c No unintended consequences found, but the document does not discuss inaccuracies or errors in the data.
**4a/b/d- The Joint Commission has been engaged in national efforts to retool measures for use with the EHR since their inception. The Health Information Technology Standards Panel (HITSP) received funding from HHS for a CMS sponsored project to retool the Joint Commission-developed Stroke (STK) measures, so that quality data could be captured directly from the EHR as a by-product of healthcare delivery. As a member of the HITSP Quality Tiger Team, The Joint Commission served as a resource in the retooling of the stroke measures. In the past year, and following public comment of the HITSP specifications, CMS convened a small workgroup to address issues with the HITSP specifications. The Joint Commission actively participated in this effort and provided extensive input to this process. The Joint Commission is currently being consulted to participate in a Quality Data Model (QDM)-based retooling effort of the stroke measures.   4c- no inadequacies noted at this time- the measure has been in effect for several years.
**Based on feedback from participating hospitals supports high

0441 STK-10: Assessed for Rehabilitation

**some issues relative to the fact that this may be abstracted from the medical record by a person other than the individual obtaining original info JCHO re-tooling EHR measure use

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-5; N-0

Rationale: Assess for rehabilitation has been a stroke core measure for the past several years. There is adequate evidence that disparities continue to exist in rehabilitation for stroke patients. Studies indicate that early evaluation of rehabilitation need and early implementation improves patient outcomes.

**High impact on most criteria

Additional Comments/Questions: this is an important metric to track to improve overall outcomes of patients who have experienced stroke however lack of state and federal funds continue to cause many patients without insurance or funding to go without the proper rehabilitation services.

Workgroup Call Summary

Importance (opportunity for improvement):

- ***Developers note a high level of performance for this measure (97.4%)—is it topped out?*** The developers noted that the literature suggests that while two-thirds of stroke patients could benefit from rehab services, less than a third actually receive rehab services—and that if a patient is never assessed, then rehab is need is never even considered. They also noted that hospitals want to meet this measure by using nursing assessments (this is not permitted: the assessment must be done by a qualified member of the rehab team).

Scientific Acceptability (reliability)

- ***What kinds of services count as rehab?*** The developer noted that the details of what counts is provided in the data element definition that is found in the alphabetical data dictionary. *NQF note: The link to this data dictionary is found in section S.2, which is located at the beginning of the Reliability and Validity section of the measure submission form. Subsequent to the workgroup call, the developer provided the alphabetical data dictionary file and this has been posted on the project SharePoint site.*

Additional staff notes

- Reliability was tested at the data element level.
- Because their data element reliability testing involved a rater whom they consider an authoritative source, their data reliability testing may also be considered validity testing at the data element level.
- Measure #2022 is, per the NQF definition, a competing measure. We will discuss measure superiority and/or harmonization during the in-person meeting.

0467 Acute Stroke Mortality Rate (IQI 17)

Status: Maintenance, Original Endorsement: Jun 23, 2008

Description: Percent of discharges with an in-hospital death among cases with a principal diagnosis code for stroke

Numerator Statement: Number of deaths among cases meeting the inclusion and exclusion rules for the denominator

Denominator Statement: All discharges, age 18 years and older, with a principal diagnosis code for stroke

Exclusions: Exclude cases:
• transferring to another short-term hospital
• MDC 14 (pregnancy, childbirth, and puerperium)
• with missing discharge disposition, gender, age, quarter, year or principal diagnosis

Adjustment/Stratification: Statistical risk model The predicted value for each case is computed using a hierarchical model (logistic regression using Generalized Estimating Equations (GEE) to account for clustering of patients within hospitals) and covariates for gender, age (in 5-year age groups pooled), APR-DRG and APR-DRG Risk of Mortality subclass, MDC and availability of Point of Origin (UB-04). The reference population used in the regression is the universe of discharges for states that participate in the HCUP State Inpatient Data (SID) for the years 2008, a database consisting of 42 states and approximately 30 million adult discharges.

Intercept	
Sex	Female
Age	18 to 59
Age	65 to 84
Age	85+
APR-DRG	´0211´
APR-DRG	´0212´
APR-DRG	´0213´
APR-DRG	´0214´
APR-DRG	´0221´
APR-DRG	´0222´
APR-DRG	´0223´ to '0224'
APR-DRG	´0231´ to '0232'
APR-DRG	´0233´
APR-DRG	´0234´
APR-DRG	´0241´
APR-DRG	´0242´
APR-DRG	´0243´
APR-DRG	´0244´
APR-DRG	´0261´ to '0263'
APR-DRG	´0264´
APR-DRG	´0441´
APR-DRG	´0442´
APR-DRG	´0443´
APR-DRG	´0444´
APR-DRG	´0452´
APR-DRG	´0453´
APR-DRG	´0454´
MDC	OTHER
NOPOUB04	UB-04 Point-of-Origin Data Not Available Not applicable

Level of Analysis: Facility

Type of Measure: Outcome

Data Source: Administrative claims

Measure Steward: Agency for Healthcare Research and Quality **Other organizations:** Battelle Memorial Institute, Stanford University and the University of California-Davis

Workgroup Evaluation Results

The following evaluation ratings and comments are from the **Committee Reviewers:** Risha Gidwani; David Knowlton; Terry Richmond; Ramon Bautista   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-3; N-0

0467 Acute Stroke Mortality Rate (IQI 17)

1a. Impact: H-2; M-1; L-0; I-0; **1b. Performance Gap:** H-2; M-1; L-0; I-0

<u>Rationale:</u> Variation in risk-adjusted mortatlity between hospital types, sized, locations.

**1a. Stroke is indeed a major cause of death, although the health goal/priority should be to decrease preventable deaths due to strokes.

1c. Evidence (*based on decision logic*): <mark>Y-3; N-0</mark>   IF a **Health Outcome**, rationale supports: <mark>Y-2; N-1; NA-0</mark>

Quantity: H-0; M-0; L-0; I-0; Quality: H-0; M-0; L-0; I-0; Consistency: H-0; M-0; L-0; I-0

<u>Rationale:</u>

---

2. Scientific Acceptability of Measure Properties (*based on decision logic*): <mark>Y-1; N-2</mark>

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

<u>Rationale:</u> Reliability Testing data section (2a2.2) showed variation. Data showing values from, for example. one-half of randomly sampled subgroup vs. other half of randomly sampled subgroup would indicate reliability of models and ratios.   Validity: Information provided for specifications of risk-adjustment model were for all AHRQ risk-adjustment models. Specific information about the stroke model, including covariates used, and beta-coefficients, is necessary. Additionally, the following information is requested: 1. How is alpha, the reference population rate, derived? 2. Please explain the statement, "If more than one POA indicators are present, the maximum value is considered" (page 10).   3. It is not clear how X is an "improved" vector of binary explanatory variables compared with Z. 4. More information about their approach to missing data for stroke would be useful. Does the approach assume that data are missing at random? If so, was this tested? 5. Do equations 5 and 6 on page 11 mean that imputed variables for X prime and P prime could be 50%? 6. Please explain more about the fitting of the B-subscript-y coefficients, specifically for stroke mortality. 7. In the Risk Adjusted Rate equation on page 16, please confirm that alpha comes from observed data (rather than predicted). 8. The section on composite measures is difficult to understand. More text explanation of the approach and rationale, including information that is specific to the stroke model, would be elucidating. 8. Page 19 - principal components and factor analysis are two different approaches.

**2a: one concern is that by definition in-hospital mortality is not standardized from a time perspective - thus systems that have seamless mechanisms to transfer patients doing poorly to long-term care may show lower in-hospital mortality.   2b. Insuffient- Uses literature to establish criterion validity - proportion of records with ICD9 codes compared to a gold standad (chart abstraction) for correct diagnosis of stroke. Sensitivity & specificity good with relatively tight 95%CI and kappas ranging from 0.72-0.88. So valdiity of denominator (stroke) is good but no information on numerator (# of stroke deaths). Addresses threats (transfer to anothe aucte care hospital) and shows not statistical differences - doesn't address other exclusions (pregnancy, missing data).

**This has to be a risk adjusted measure because the profile of stroke patients across different healthcare systems can vary widely.

---

3. Usability: <mark>H-1; M-0; L-1; I-1</mark>

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

<u>Rationale:</u> More information about risk-adjustment approach are needed before assessing usability of measure.

**3a - 18 states/systems publicly report, although not all states report this specific outcome (e.g., Oregon indicates data not available for this measure). Used by others as well.

**The measure has to be risk adjusted to be useful for public consumption and by itself does not necessarily reflect on the quality of care a patient received.

---

4. Feasibility: <mark>H-2; M-0; L-1; I-0</mark>

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

<u>Rationale:</u>

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-0; N-3</mark>

<u>Rationale:</u> More information about risk-adjustment is needed -- the methods document provided is a generic document for all AHRQ indicators. A document specific to stroke is needed. Additionally, the specific model for stroke is not shown, and this is necessary in order for the NQF steering committee to evaluate the face validity of predictor variables and their beta-coefficients (request that beta-coefficients be presented in terms of probabilities rather than ORs for ease of interpretation)>

**Overall, I think the answer is yes, but insufficient information for 2b (validity) would be helpful for a final decision.

**The measure has to be risk adjusted to allow for meaningful comparison across healthcare organizations.

<u>Additional Comments/Questions</u>:

Workgroup Call Summary

Scientific acceptability

- *Risk adjustment methodology:*
  - **Was specific risk-adjustment model information provided?** Yes, this was provided by the developer (see Table 9 on page 11 in the document titled: *0467 - Risk Adjustment Tables.pdf*).
  - **Is stroke severity captured in the risk model?** The score from the NIHSS is not included in the risk-adjustment model because it is not available in claims data.
  - **APR-DRG categories in the risk model**: Several workgroup members requested additional information about what

0467 Acute Stroke Mortality Rate (IQI 17)

the APR-DRG categories mean. One member noted that this information might provide insight on whether other measures of stroke severity or co-morbidities are included in the model.

- *Technical questions about the risk-adjustment model*: One workgroup member posed several very technical questions related to the risk-adjustment methodology. These questions have been forwarded to the developer and the responses will be shared with the full Steering Committee prior to the in-person meeting. In addition, an NQF staff review of measure testing—including the risk-adjustment methodology—will be provided.

- *Data element validity*: One workgroup member noted that the results from testing of data element validity did not include information about the numerator (i.e., using the claims data to identify death).

- *Exclusions:* One workgroup member noted that the number of records excluded from the measure (because of pregnancy or because of missing data) was not provided.

Other points of discussion
- One workgroup member noted in their comments that systems with seamless mechanisms to transfer patients to long-term care may show lower in-hospital mortality.

2026 Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following an acute ischemic stroke hospitalization

Status: New Submission

Description: The measure estimates a hospital-level, risk-standardized mortality rate (RSMR) for patients discharged from the hospital with a principal diagnosis of acute ischemic stroke. Mortality is defined as death from any cause within 30 days of the index admission date for patients discharged from the hospital with a principal diagnosis of acute ischemic stroke.

Numerator Statement: The outcome for this measure is 30-day all-cause mortality. We define mortality as death from any cause within 30 days from the index admission date for patients discharged from the index hospital with a principal diagnosis of acute ischemic stroke.

Denominator Statement: The cohort includes admissions for patients age 65 years or older discharged from the hospital with a principal diagnosis of acute ischemic stroke (ICD-9-CM codes 433.x1, 434.x1, 436) and with a complete claims history for the 12 months prior to admission.

Exclusions: An index admission is the hospitalization considered for mortality outcome.

The measure excludes admissions for patients:

• transferred from another acute care hospital (because the death is attributed to the hospital where the patient was initially admitted);

• with inconsistent or unknown mortality status or other unreliable data (e.g. date of death precedes admission date).

• who were discharged alive and against medical advice (AMA) (because providers did not have the opportunity to deliver full care and prepare the patient for discharge);

• enrolled in the Medicare Hospice program any time in the 12 months prior to the index hospitalization including the first day of the index admission (since it is likely these patients are continuing to seek comfort measures only).

Adjustment/Stratification: Statistical risk model Our approach to risk adjustment was tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes".1

The measure employs a hierarchical logistic regression model to create a hospital-level 30-day RSMR. In brief, the approach simultaneously models two levels (patient and hospital) to account for the variance in patient outcomes within and between hospitals(Normand & Shahian, 2007). At the patient level, each model adjusts the log-odds of mortality within 30 days of admission for age and selected clinical covariates. The second level models the hospital-specific intercepts as arising from a normal distribution. The hospital intercept represents the underlying risk of mortality, after accounting for patient risk. See section 2a1.20. Calculation Algorithm/Measure Logic for more detail.

Candidate and Final Risk-adjustment Variables: The measure was initially developed using Medicare FFS 2007 claims data. Candidate variables were patient-level risk adjustors that were expected to be predictive of mortality, based on empirical analysis, prior literature, and clinical judgment, including age and indicators of comorbidity and disease severity. For each patient, covariates are obtained from Medicare claims extending 12 months prior to and including the index admission. The model adjusts for case mix differences based on the clinical status of patients at the time of admission. We used condition categories (CCs), which are clinically meaningful groupings of more than 15,000 ICD-9-CM diagnosis codes, and combinations of CCs as candidate variables. A file which contains a list of the ICD-9-CM codes and their groupings into CCs is available on www.qualitynet.org (http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1182785083979)

We did not risk-adjust for CCs that were possible adverse events of care and that were only recorded in the index admission. Only comorbidities that conveyed information about the patient at that time or in the 12 months prior, and not complications that arose during the course of the hospitalization were included in the risk-adjustment.

Following initial model development, in response to suggestions from our working group and Technical Expert Panel (TEP) members, we evaluated the mortality rates of patients admitted for stroke after having been evaluated at a different hospital's emergency department. Our experts expressed concern that such patients may be at higher risk and that the admitting hospital would not have had the opportunity to evaluate and treat such patients at first presentation. They also felt that certain hospitals may receive substantially greater proportions of patients transferred from outside EDs. Based on our analyses, we updated the measure to include a risk factor that indicates if a patient was transferred in from an outside ED, that is, the patient was seen in a different hospital's ED prior to being admitted for the index admission. This revision was done using 2008 data.

Frequencies and odds ratios for the model are presented below (2008 Medicare FFS patients aged 65 and older; n=175,267 admissions): Final set of risk-adjustment variables:

Variable//Frequency (%)//Odds Ratio (95% confidence interval)

• Transfer from another ED/Frequency= 5.64/OR (95% CI)= 1.37 (1.29-1.45)

Demographic

• Age-65 (continuous)/mean (SD)=15.31 (7.93)/OR (95% CI)= 1.069 (1.067-1.07)

• Male /Frequency= 40.28/OR (95% CI)= 0.99 (0.96-1.03)

Cardiovascular/Cerebrovascular

• Congestive Heart Failure /Frequency= 26.03/OR (95% CI)= 1.38 (1.34-1.43)

• Valvular and Rheumatic Heart Disease /Frequency= 23.03/OR (95% CI)= 0.87 (0.84-0.89)

• Congenital Cardiac/Circulatory Defects /Frequency= 2.04/OR (95% CI)= 0.71 (0.64-0.8)

• Hypertensive Heart Disease /Frequency= 6.54/OR (95% CI)= 0.83 (0.78-0.88)

2026 Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following an acute ischemic stroke hospitalization

- Specified Heart Arrhythmias /Frequency= 29.37/OR (95% CI)= 1.59 (1.54-1.64)
- Cerebral Hemorrhage /Frequency= 1.88/OR (95% CI)= 1.16 (1.06-1.27)
- Ischemic or Unspecified Stroke /Frequency= 24.81/OR (95% CI)= 1.00 (0.96-1.03)
- Precerebral Arterial Occlusion and Transient Cerebral Ischemia /Frequency= 22.83/OR (95% CI)= 0.82 (0.8-0.85)
- Cerebral Atherosclerosis and Aneurysm /Frequency= 10.67/OR (95% CI)= 0.83 (0.80-0.87)
- Hemiplegia/Hemiparesis /Frequency= 5.60/OR (95% CI)= 1.17 (1.10-1.24)

Comorbidities
- History of Infection/Frequency= 26.72/OR (95% CI)= 1.15 (1.11-1.18)
- Metastatic Cancer and Acute Leukemia and Other Major Cancers /Frequency= 3.65/OR (95% CI)= 2.77 (2.61-2.95)
- Lymphatic, Head and Neck, Brain, Breast, Colorectal and Other Major Cancers/Frequency= 23.92/OR (95% CI)= 0.92 (0.89-0.95)
- Protein-Calorie Malnutrition /Frequency= 5.42/OR (95% CI)= 1.69 (1.61-1.77)
- Other Significant Endocrine and Metabolic Disorders /Frequency= 75.98/OR (95% CI)= 0.75 (0.72-0.77)
- Other Gastrointestinal Disorders /Frequency= 43.64/OR (95% CI)= 0.90 (0.88-0.93)
- Disorders of the Vertebrae and Spinal Discs /Frequency= 17.06/OR (95% CI)= 0.89 (0.86-0.93)
- Osteoarthritis of Hip or Knee /Frequency= 10.36/OR (95% CI)= 0.82 (0.78-0.86)
- Other Musculoskeletal and Connective Tissue Disorders /Frequency= 63.50/OR (95% CI)= 0.86 (0.84-0.89)
- Iron Deficiency and Other/Unspecified Anemia and Blood Disease /Frequency= 31.86/OR (95% CI)= 1.09 (1.05-1.12)
- Dementia or senility /Frequency= 28.64/OR (95% CI)= 1.24 (1.20-1.28)
- Major Psychiatric Disorders /Frequency= 9.12/OR (95% CI)= 1.08 (1.04-1.13)
- Quadriplegia, Other Extensive Paralysis /Frequency= 1.54/OR (95% CI)= 1.39 (1.26-1.53)
- Multiple Sclerosis /Frequency= 10.27/OR (95% CI)= 0.83 (0.79-0.87)
- Seizure Disorders and Convulsions /Frequency= 6.92/OR (95% CI)= 1.27 (1.21-1.33)
- Hypertension /Frequency= 88.00/OR (95% CI)= 0.77 (0.74-0.81)
- Peripheral Vascular Disease /Frequency= 23.02/OR (95% CI)= 1.07 (1.04-1.11)
- Chronic Obstructive Pulmonary Disease /Frequency= 21.92/OR (95% CI)= 1.06 (1.03-1.10)
- Pneumonia /Frequency= 17.36/OR (95% CI)= 1.49 (1.44-1.54)
- Pleural Effusion/Pneumothorax /Frequency= 6.92/OR (95% CI)= 1.13 (1.07-1.18)
- Other Eye Disorders /Frequency= 19.34/OR (95% CI)= 0.91 (0.88-0.94)
- Other Ear, Nose, Throat, and Mouth Disorders /Frequency= 26.99/OR (95% CI)= 0.87 (0.84-0.90)
- Dialysis Status /Frequency= 1.47/OR (95% CI)= 1.38 (1.24-1.52)
- Renal Failure /Frequency= 15.45/OR (95% CI)= 1.16 (1.12-1.21)
- Urinary Tract Infection /Frequency= 21.55/OR (95% CI)= 1.14 (1.10-1.18)
- Male Genital Disorders /Frequency= 11.95/OR (95% CI)= 0.78 (0.74-0.82)
- Decubitus Ulcer of Skin /Frequency= 2.52/OR (95% CI)= 1.29 (1.20-1.39)
- Chronic Ulcer of Skin, Except Decubitus /Frequency= 5.52/OR (95% CI)= 1.16 (1.10-1.23)
- Other Dermatological Disorders /Frequency= 29.38/OR (95% CI)= 0.92 (0.89-0.95)

References:

1. Krumholz HM, Brindis RG, Brush JE, et al. 2006. Standards for Statistical Models Used for Public Reporting of Health Outcomes: An American Heart Association Scientific Statement From the Quality of Care and Outcomes Research Interdisciplinary Writing Group: Cosponsored by the Council on Epidemiology and Prevention and the Stroke Council Endorsed by the American College of Cardiology Foundation. Circulation 113: 456-462.

2. Normand S-LT, Shahian DM. 2007. Statistical and Clinical Aspects of Hospital Outcomes Profiling. Stat Sci 22 (2): 206-226. N/A

Level of Analysis: Facility

Type of Measure: Outcome

Data Source: Administrative claims, Other

Measure Steward: Centers for Medicare & Medicaid Services (CMS) **Other organizations:** MPR: Mathematica Policy Research; RTI: Research Triangle Institute

Workgroup Evaluation Results

The following evaluation ratings and comments are from the **Committee Reviewers:** Terry Richmond; Risha Gidwani; Ramon Bautista *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-3; N-0
1a. Impact: H-3; M-0; L-0; I-0; 1b. Performance Gap: H-3; M-0; L-0; I-0
Rationale:
1c. Evidence (*based on decision logic*): Y-3; N-0   IF a **Health Outcome**, rationale supports: Y-2; N-1; NA-0
Quantity: H-0; M-0; L-0; I-0; Quality: H-0; M-0; L-0; I-0; Consistency: H-0; M-0; L-0; I-0

2026 Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following an acute ischemic stroke hospitalization
<u>Rationale</u>:
2. Scientific Acceptability of Measure Properties (*based on decision logic*): <mark>Y-1; N-2</mark> **2a. Reliability:** H-0; M-2; L-1; I-0; **2b. Validity:** H-1; M-0; L-0; I-2 <u>Rationale</u>: Compares medical record abstraction with administrative data which includes both numberator and denominator information. c stat 0.8. I would reserve final judgment on this criterion until we have additional information on inclusion ot stroke severity in the risk adjustment model. **Reliability statistics showed the agreement between the RSMR for each hospital via the administrative dataset was 0.4, which is moderate. Validity: 1. The measure itself as submitted is the predicted-to-expected mortalities. Why not look at observed-to-expected mortalities? In this predicted-to-observed model, both the numerator and denominator are based on prediction models – why are actual data on mortalities not driving the numerator of the ratio? 2. What is the rationale for multiplying the ratio of predicted-to-observed mortalities by the national unadjusted mortality rate?1. How was the interval estimate of s(i) created through boostrapping (page 17 of methods report)? 3. Table 4 of methods report shows many variables are not linearly related to Y. 4. What are the p-values associated with the chi-squared test in Table 6 in the methods report? (The test statistic is quite large but the level of significance is needed.) 5. What family of distributions was used for the generalized linear model? 6 Why is ischemic stroke being included as a covariate if all patients in the model have ischemic stroke? 7. Many covariates listed on page 10 of measure specifications report have OR <1.0 that are of potential concern. For example, a patient with hypertension would have lower odds of mortality than a patient without hypertension. Is this because the dummy variable was specified as "having hypertension = 0"?
3. Usability: <mark>H-1; M-1; L-0; I-1</mark> (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*) <u>Rationale</u>: not currently used in QI but could be useful **The questions about validity need to be settled before assessing these question.
4. Feasibility: <mark>H-1; M-2; L-0; I-0</mark> (*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*) <u>Rationale</u>: measure not in operational use but all elements part of electronic health record **Required data elements (i.e. mortality) do not seem to be routinely gathered nor is there a data collection strategy in place.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-1; N-2</mark> <u>Rationale</u>: I would like further infromation and discussion about the presence or absence of stroke severity as part of risk adjustment prior to supporting endorsement. **This is a preliminary conclusion -- the questions raised about validity will first need to be answered before coming to a final conclusion. **Despite current feasibility limitations, this measure should be implemented and can serve as one index of quality of stroke care. <u>Additional Comments/Questions</u>:
Workgroup Call Summary

Scientific acceptability
- **Risk adjustment methodology:**
  - **Why is predicted mortality used in the numerator rather than observed mortality?** Developers noted that using predicted mortality allows them to adjust for hospitals that have a very low sample size. Developers will provide additional an response regarding this question prior to the in-person meeting. In addition, during the in-person meeting, NQF staff will provide some background on this question.
  - **Technical questions about the risk-adjustment model**: One workgroup member posed several very technical questions related to the risk-adjustment methodology. These questions have been forwarded to the developer and the responses will be shared with the full Steering Committee prior to the in-person meeting.

Usability
- One workgroup member noted that developers did not discuss plans for how this measure would be used in public reporting/quality improvement programs.

Other points of discussion
- Developers noted that the current measure is currently specified only for patients 65 and older; however, they are planning to test the measure in younger age groups. Once that testing is complete they will bring the newly-specified measure back to NQF for a maintenance review.

2027 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following an acute ischemic stroke hospitalization

Status: New Submission

Description: The measure estimates a hospital-level risk-standardized readmission rate (RSRR) for patients discharged from the hospital with a principal diagnosis of acute ischemic stroke. We define this as readmission for any cause within 30 days from the date of discharge of the index stroke admission.

Numerator Statement: The outcome for this measure is 30-day all-cause readmission. We define all-cause readmission as readmission for any cause within 30 days from the date of discharge of the index stroke for patients discharged from the hospital with a principal diagnosis of ischemic stroke. If a patient has one or more admissions (for any reason) within 30 days after discharge from the index admission, only one is counted as a readmission.

Denominator Statement: The cohort includes admissions for patients age 65 years or older discharged from the hospital with a principal diagnosis of ischemic stroke (ICD-9-CM codes 433.x1, 434.x1, 436) and with a complete claims history for the 12 months prior to admission.

Exclusions: An index admission is the hospitalization considered for the readmission outcome (readmitted within 30 days of the date of discharge from the initial admission).

The measure excludes admissions for patients:

• with an in hospital death (because they are not eligible for readmission).

• transferred to another acute care facility (because the readmission is attributed to the hospital that discharges the patient to a non-acute setting).

• discharged alive and against medical advice (AMA) (because providers did not have the opportunity to deliver full care and prepare the patient for discharge).

• without at least 30 days post-discharge claims data (because the 30-day readmission outcome cannot be assessed in this group).

In addition, if a patient has more than one admission within 30 days of discharge from the index admission, only one is counted as a readmission, as we are interested in a dichotomous yes/no readmission outcome, as opposed to the number of readmissions. No admissions within 30 days of discharge from an index admission are considered as additional index admissions, thus no hospitalization will be counted as both a readmission and an index admission. The next eligible index admission is 30 days after the discharge date of the previous index admission.

Adjustment/Stratification: Statistical risk model Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes"1.

The measure employs a hierarchical logistic regression model to create a hospital-level 30-day RSRR. This approach to modeling appropriately accounts for the structure of the data (patients clustered within hospitals), the underlying risk due to patients' comorbidities, and sample size at a given hospital when estimating hospital readmission rates. In brief, the approach simultaneously models two levels (patient and hospital) to account for the variance in patient outcomes within and between hospitals.2 At the patient level, the model adjusts the log-odds of readmission within 30 days of discharge for age and selected clinical covariates. The second level models hospital-specific intercepts as arising from a normal distribution. The hospital-specific intercepts represent the hospital contribution to the risk of readmission, after accounting for patient risk and sample size, and can be inferred as a measure of quality. The hospital-specific intercepts are given a distribution in order to account for the clustering (non-independence) of patients within the same hospital. If there were no differences among hospitals, then after adjusting for patient risk, the hospital intercepts should be identical across all hospitals.

Candidate and Final Risk-adjustment Variables: The measure was developed using Medicare FFS 2007 claims data. Candidate variables were patient-level risk-adjustors that were expected to be predictive of readmission, based on empirical analysis, prior literature, and clinical judgment, including age and indicators of comorbidity and disease severity. For each patient, covariates are obtained from Medicare claims extending 12 months prior to and including the index admission. The model adjusts for case mix differences based on the clinical status of patients at the time of admission. We used condition categories (CCs), which are clinically meaningful groupings of more than 15,000 ICD-9-CM diagnosis codes, and combinations of CCs as candidate variables. A file which contains a list of the ICD-9-CM codes and their groupings into CCs is available on

http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1182785083979). We did not risk-adjust for CCs that were possible adverse events of care and that were only recorded in the index admission. Only comorbidities that conveyed information about the patient at that time or in the 12 months prior, and not complications that arose during the course of the hospitalization were included in the risk-adjustment.

Frequencies and odds ratios for the 2007 cohort (n=174,024 admissions) are presented below.

Final set of risk-adjustment variables:

Variable//Frequency (%)//Odds Ratio (95% confidence interval)

Demographic

•   Age-65 (continuous)/Mean (SD)=80.12(7.83)/ OR (95% CI)=1.004(1.003 - 1.006)

•   Male/Frequency =40.44/ OR (95% CI)=1.045(1.016 - 1.045)

Cardiovascular/Cerebrovascular

2027 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following an acute ischemic stroke hospitalization

- Congestive Heart Failure (CC 80)/Frequency =25.68/ OR (95% CI)=1.221(1.182 - 1.261)
- Hypertensive heart disease (CC 90)/Frequency =6.91/ OR (95% CI)=1.100(1.047 - 1.157)
- Cerebral Hemorrhage (CC 95)/Frequency =1.81/ OR (95% CI)=1.079(0.954 - 1.182)
- Ischemic or Unspecified Stroke (CC 96)/Frequency =26.41/ OR (95% CI)=1.042(1.008 - 1.078)
- Cerebrovascular Disease (CC 97)/Frequency =23.75/ OR (95% CI)=1.045(1.010 - 1.080)
- Hemiplegia, paraplegia, paralysis, functional disability (CC 100-102)/Frequency =9.70/ OR (95% CI)=0.951(0.907 - 0.997)
- Vascular or circulatory disease (CC 104-106)/Frequency =31.09/ OR (95% CI)=1.070(1.038 - 1.103)

Comorbid Conditions
- Metastatic cancer and acute leukemia (CC 7)/Frequency =2.27/ OR (95% CI)=1.264(1.163 - 1.373)
- Cancer (CC 8-12)/Frequency =18.52/ OR (95% CI)=1.034(0.998 - 1.071)
- Diabetes and DM complications (CC 15-20, 119-120)/Frequency =37.84/ OR (95% CI)=1.156(1.124 - 1.364)
- Protein-calorie malnutrition (CC 21)/Frequency =4.45/ OR (95% CI)=1.288(1.216 - 1.364)
- Disorders of Fluid/Electrolyte/Acid-Base (CC 22-23)/Frequency = 23.72/ OR (95% CI)=1.142(1.104 - 1.181)
- Obesity/disorders of thyroid, cholesterol, lipids (CC 24)/Frequency = 68.03/ OR (95% CI)=0.916(0.890 - 0.943)
- Severe Hematological Disorders (CC 44)/Frequency = 1.53/ OR (95% CI)=1.266(1.153 - 1.391)
- Iron Deficiency and Other/Unspecified Anemias and Blood Disease (CC 47)/Frequency = 30.90/ OR (95% CI)=1.142(1.108 - 1.178)
- Dementia and senility (CC 49-50)/Frequency = 28.56/ OR (95% CI)=1.015(0.985 - 1.047)
- Quadriplegia, paraplegia, functional disability (CC 67-69, 177-178)/Frequency = 1.99/ OR (95% CI)=1.139(1.046 - 1.242)
- Seizure Disorders and Convulsions (CC 74)/Frequency = 7.45/ OR (95% CI)=1.161(1.107 - 1.218)
- COPD (CC 108)/Frequency =22.96/ OR (95% CI)=1.133(1.098 - 1.170)
- Other lung disorder (CC 115)/Frequency =22.04/ OR (95% CI)=1.082(1.047 - 1.117)
- End-stage renal disease or dialysis (CC 130)/Frequency =1.51/ OR (95% CI)=1.356(1.237 - 1.487)
- Renal Failure (CC 131)/Frequency =14.29/ OR (95% CI)=1.163(1.117 - 1.211)
- Other urinary tract disorders (CC 136)/Frequency =18.57/ OR (95% CI)=1.101(1.064 - 1.140)
- Decubitus ulcer or chronic skin ulcer (CC 148-149)/Frequency =6.79/ OR (95% CI)=1.079(1.026 - 1.134)
- Major Symptoms, Abnormalities (CC 166)/Frequency =61.63/ OR (95% CI)=1.098(1.063 - 1.134)

References:

1. Krumholz HM, Brindis RG, Brush JE, et al. 2006. Standards for Statistical Models Used for Public Reporting of Health Outcomes: An American Heart Association Scientific Statement From the Quality of Care and Outcomes Research Interdisciplinary Writing Group: Cosponsored by the Council on Epidemiology and Prevention and the Stroke Council Endorsed by the American College of Cardiology Foundation. Circulation 113: 456-462.
2. Normand S-LT, Shahian DM. 2007. Statistical and Clinical Aspects of Hospital Outcomes Profiling. Stat Sci 22 (2): 206-226. N/A

Level of Analysis: Facility

Type of Measure: Outcome

Data Source: Administrative claims

Measure Steward: Centers for Medicare & Medicaid Services (CMS) **Other organizations:** MPR: Mathematica Policy Research; RTI: Research Triangle Institute

Workgroup Evaluation Results

The following evaluation ratings and comments are from the **Committee Reviewers:** Terry Richmond; Risha Gidwani; Ramon Bautista
(comments separated by asterisks)

Importance to Measure and Report (*based on decision logic*): Y-3; N-0
1a. Impact: H-2; M-1; L-0; I-0; 1b. Performance Gap: H-2; M-1; L-0; I-0
Rationale:
1c. Evidence (*based on decision logic*): Y-3; N-0   IF a **Health Outcome**, rationale supports: Y-3; N-0; NA-0
Quantity: H-0; M-0; L-0; I-0; Quality: H-0; M-0; L-0; I-0; Consistency: H-0; M-0; L-0; I-0
Rationale:

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-1; N-2
2a. Reliability: H-1; M-2; L-0; I-0; 2b. Validity: H-0; M-1; L-0; I-2
Rationale: 2b) More information and discussion about the inclusion of stroke severity in the risk adjustment is warranted. Plan for publuc reporting and discriminating performance is not developed.
**Validity: the following questions remain after reading the methodology report 1. What is the rational for including cancer as a covariate when it did not meet the bootstrap threshold and the Confidence Interval for the Odds Ratio includes 1.0? (page 20) 2. Why is "ischemic or unspecified stroke" included as a covariate when all patients should have had ischemic stroke, according to the measure specifications?   3.   The measure itself as submitted is the predicted-to-expected readmissions. Why not look at observed-to-expected readmissions? In this predicted-to-observed model, both the numerator and denominator are based on prediction models – why are actual data on

2027 Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following an acute ischemic stroke hospitalization

readmissions not driving the numerator of the ratio? 4.   What is the rationale for multiplying the ratio of predicted-to-observed readmissions by the national unadjusted readmission rate? 5.   What family of distributions was used for the generalized linear model (page 25)? 6.   The area under the ROC curve for the model is 0.602 (page 28). A value of 0.50 indicates that the model has no ability to predict readmissions. What is the authors' rationale for using a prediction model with a low value of 0.602? a.   The authors do a nice job of showing reliability of the model (page 31), but the validity of the model is still a concern. 7.   Table 7 on page 33 and Table 14 on page 43 should have p-values included along with the chi-squared statistics.

3. Usability: H-1; M-1; L-0; I-1
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
<u>Rationale:</u> More discussion of how to interpret a predicted-to-expected value is needed.

4. Feasibility: H-2; M-1; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
<u>Rationale:</u>

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-1; N-2
<u>Rationale:</u> Would like further discussion on the inclusion or absence of stroke severity in the risk adjustment and the implications of this prior to considering endorsement
**This is a preliminary conclusion; more details on the modeling process and rationale are needed before coming to a final conclusion.
<u>Additional Comments/Questions</u>: Consider excluding elective readmissions from the denominator.

Workgroup Call Summary

Scientific acceptability
- *Risk adjustment methodology:*
  - **Why is the area under the ROC curve so low (c=.602)?** Developers noted that they only adjust for factors that are present at the time of admission—but that patient co-morbidity factors are not as strong in predicting readmission as they are in predicting mortality. They believe that readmissions truly reflect hospital quality (particularly transitions and follow-up care)—and quality-related factors should <u>not</u> be included in risk-adjustment models.
  - **Has there been any thought about including other types of data in the risk-adjustment model (e.g., ambulance pattern, distance from hospital, etc.)?**   Developers acknowledged that this is an interesting question, but noted that they have not conducted analyses to include data from other (non-claims) sources.
  - **Technical questions about the risk-adjustment model***: One workgroup member posed several very technical questions related to the risk-adjustment methodology. These questions have been forwarded to the developer and the responses will be shared with the full Steering Committee prior to the in-person meeting.
- **Are elective admissions included in this measure?** Developers noted that admissions for several procedures that might be scheduled within 30 days of an ischemic stroke hospitalization have been excluded from the measure. ***NOTE***: *NQF staff have asked the developers to discuss if/how the exclusion of planned readmissions in this measure differs from the planned readmissions that are excluded in other CMS readmission measures that are endorsed by NQF.*

Other points of discussion
- During public comment, representatives from the American Academy of Neurology noted several concerns with the measure. These comments are included in a letter that the AAN submitted to CMS in 2010. This letter has been made available to the full Steering Committee (posted to the SharePoint site). Developers for this measure plan to provide a written response to at least some of the concerns in this letter prior to the in-person meeting.

0442 Functional Communication Measure: Writing

Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: Number of stroke patients who make progress as defined by an increase of one or more levels on the Writing Functional Communication Measure (FCM).
Numerator Statement: Number of stroke patients who make progress on the Writing Functional Communication Measure.
Denominator Statement: Number of stroke patients scored on the Writing Functional Communication Measure.
Exclusions: Patients discharged from speech-language pathology services after only one treatment session.
Patients who are not candidates for memory treatment as demonstrated by the highest level of functioning on admission (Level 7 on the Writing Functional Communication Measure).
Patients using an augmentative-alternative communication system.
Adjustment/Stratification: Stratification by risk category/subgroup   N/A.
Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System
Type of Measure: Outcome
Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records
Measure Steward: American Speech-Language-Hearing Association

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Mary Van de Kamp; Jordan Eisenstock; Jane Sullivan; Jane Sullivan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-4; N-0
1a. Impact: H-3; M-1; L-0; I-0; **1b. Performance Gap:** H-3; M-1; L-0; I-0
Rationale: 1a. Developer states that in 2011, 10% (1527) of patients receiving SLP services were treated for a writing disorder.   1b. Of the patients treated for this disorder, 30.3% failed to make progress. This seems like an undesirably high percentage. Developer also identifies disparities i score increases by race, insurance, baseline status, age & diagnosis. Further study of these disparities is waranted.
**Numerator, denominator, exclusions clear. Evidence for impact based on stroke prevalence, mortality & morbidity.   Evidence for benefits of multi-disciplinary stroke care. HC organizations that track this measure have seen improving rates over time. Rates from JCHO and PCDAS demonstrate high adherence but a slight performance gap (3-4%) Disparity in stroke care/outcomes by race, gender, & SES (some disparities are significant)
1c. Evidence (*based on decision logic*): Y-4; N-0   IF a **Health Outcome**, rationale supports: Y-3; N-0; NA-0
Quantity: H-2; M-0; L-0; I-0; **Quality:** H-1; M-1; L-0; I-0; **Consistency:** H-1; M-1; L-0; I-0
Rationale: Quantity - EBRSR cites 2000 studies including 1078 RCTs & 5 metanalyses on the effectiveness of stroke rehab. + findings include: increased function, reduced mortality & institutional care. Quality/Consistency - inconsistencies in design, methodology make consensus   difficult   No evidence of harm   Cost effectiveness is uncertain   EBRSR & AHA/ASA CPGs both strongly recommend

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-0
2a. Reliability: H-2; M-2; L-0; I-0; **2b. Validity:** H-1; M-3; L-0; I-0
Rationale: Numerator, denominator,exclusions, statistical risk model, data source clear. Time window ambiguous (? important). Data collection from multiple sources (but clinically realistic). Reliability testing based on hypothetical vignettes vs real clinical use. Validity testing via correlation coefficients btw treated patients' satisfaction and clinicians' scores that was reported to find some evidence of convergent, discriminant and construct validity. Is client satisfaction the appropriate comparator to establish validity???
**Numerator, denominator, exclusions, target population clear. Sampling methodology, data analysis clear. JCHO data for 2010-1 show high agreement rate (98.3%)   Measure has been in use since 2009 Validity initially assessed via survey & focus groups of hospitals; since implementation validity assessed via analysis of user feedback (72 submissions) and expert technical advisory panel   some concern about exclusions relative to attainment of 100%   Voluntary data collection may result in overstating of % (relative to real performance)

3. Usability: H-3; M-1; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: Measure developed systematically with input from multiple sources, systematically modified in response to feedback, peer reviewed and field tested...although reliability testing was done with hypothetical cases...doesn't confirm that measure has real-world usability. Measure use is recommended by oversight entities.
**Rates included in Quality Check Website, PCNASR, EHR incentive program for hospitals & CAHs   Will become part of CMS Inpatient Quality Reporting   Measure specifications standardized & regularly updated; JCHO has process for fb   Data suggests improvements over time

4. Feasibility: H-2; M-2; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: The required data elements would be present for SLPs who have done the training and agreed to collect the data.

0442 Functional Communication Measure: Writing

**some issues relative to the fact that this may be abstracted from the medical record by a person other than the individual obtaining original info JCHO re-tooling EHR measure use

Preliminary Assessment of Criteria Met/Suitable for Endorsement: <mark>Y-4; N-0</mark>
Rationale:
Additional Comments/Questions:

Workgroup Call Summary

These notes apply to all 8 ASLHA measures

- *Initial clarifications.* Prior to discussing the individual measures, NQF staff asked the developer to respond to several questions that were relevant to all of the ASLHA functional communication measures; questions and responses are summarized below:
  - The numerator in your measure is the number of stroke patients "who make progress" in a particular area; **how is progress defined**? Movement from one level on a scale to one or more higher levels on that scale.
  - **What is the time frame** for measuring progress? Admission to a particular speech-language-pathology case load (e.g., treatment for a writing deficit) to discharge from that speech-language-pathology case load.
  - Those with only one treatment session were noted as exclusions to the measures; **how many patients have only one treatment session**? The way these measures are most commonly used is through a registry maintained by ASLHA. Those using this registry report only for those who have at least 2 treatments; therefore, we do not have any way of knowing how many have only one treatment.
  - It seems that the **denominator should include** only those patients who fall into levels one through six at admission: **is this correct**? Yes. If, at admission, they are already at the highest level of functioning, they are not candidates for treatment for that particular disorder, and would not be scored on that measure. NOTE: NQF staff have suggested that the developers modify their submission to show only six levels in the denominator details section (2a1.7).
  - **What is your risk adjustment strategy: statistical risk adjustment or stratification**? Stratification (although results from regression analysis helped to inform the stratification schemes).
  - **Should the supporting document entitled "NQFMeasureSpecifications.pdf" be dis-regarded?** Yes. This document reflected the 2008 measure specifications. The specifications have been refined since then, and the new specifications will be included in the measure submission forms.

- *Importance*
  - *Does the information presented represent a performance gap/opportunity for improvement?* If there is opportunity for improvement for a measure, then there is either variability in performance across the entities being measured or there is overall poor performance on the measure (or both), and/or variability in performance across population subgroups (i.e., disparities in performance). The developer often notes the percentage of patients who failed to make progress—and this may be interpreted as overall poor performance. However, to demonstrate *variability* in performance, generally distributional statistics (e.g., mean, percentile distributions, standard deviations, etc.) for the measure (as specified) are expected (presenting scores for randomly selected patients/facilities does not really demonstrate the variability in the distributions of the measure scores). Distributional statistics should be provided for the measure as specified—so, for example, if a measure is specified with 12 strata, distributional statistics for the 12 strata would be expected.
  - *Structure-process-outcome relationship:* Workgroup members requested additional information on this item from the developer.

- *Scientific Acceptability (validity)*
  - *Initial validity testing:* One workgroup member noted that the validity testing was done when the scales were first developed and asked if additional validity testing had been done since then. The developer stated that additional validity testing has not been done.
  - *Validity testing using vignettes*: One workgroup member noted that the weakness in using vignettes to test validity is that, in a vignette, the information about the hypothetical patient is stated up-front, whereas in real life, the provider must extract that information from the patient. Another workgroup member noted

0442 Functional Communication Measure: Writing

that this type of training is standard in the rehab setting. The developer noted that to be qualified to submit data to the ASLHA registry, providers must pass the vignette test with an 80%. Those that score less than 80% are instructed to review the training materials again and then take another test (that uses different vignettes)—but are only given 2 opportunities to pass the test. The initial pass rate is approximately 90%.

- **Validity and potential for positive bias**: The developer agreed that positive bias (up-scoring a patient's level at discharge) is a concern and discussed two avenues they are currently pursuing to discourage it (audit mechanisms and patient-reported outcomes).

-

- **Usability**
  - **Usefulness for quality improvement:** The developer did not initially include information about how this measure is/would be useful for quality improvement efforts; they will modify their submission to include this information.
  - **Reporting rates**: The developer clarified that currently, because these measures require input of data into a registry, few speech-language pathologists report on these measures through PQRS,
  - **Meaningful and understandable:** One workgroup member noted a lack of a demonstration that the results of the measures are meaningful and understandable. The developer—as part of the modifications they will be making to their submissions—will try to include specific examples of how the data have been found to be useful.

- **Additional notes**
  - **NQF staff clarification:** When reliability/validity testing is done only at the data element level **OR** only at the measure score level (but not both), the highest rating for reliability/validity that a measure is eligible for is moderate. But even then, a moderate rating should be given only if the specifications are precise and the testing results are reasonable; similarly, a rating of moderate for validity should be given only if the specifications are consistent with the evidence, if threats to validity have been addressed (including—since this is an outcome measure—adequate risk adjustment), and if testing results are reasonable.
  - **Modifications to the submission forms:** NQF staff asked the developers to modify their submissions to address many of these concerns.

0443 Functional Communicaton Measure: Swallowing
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: This measure describes the change in functional communication status subsequent to speech-language pathology treatment of patients who exhibit difficuty in swallowing.
Numerator Statement: Number of stroke patients who make progress as defined by an increase of one or more levels on the Swallowing Functional Communication Measure (FCM).
Denominator Statement: Number of stroke patients scored on the Swallowing Functional Communication Measure.
Exclusions: Patients discharged from speech-language pathology services after only one treatment session. Patients who are not candidates for memory treatment as demonstrated by the highest level of functioning on admission (Level 7 on the Swallowing Functional Communication Measure).
Adjustment/Stratification: Stratification by risk category/subgroup   N/A
Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System
Type of Measure: Outcome
Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records
Measure Steward: American Speech-Language-Hearing Association
Workgroup Evaluation Results
The following evaluation ratings and comments are from the Committee Reviewers: Mary Van de Kamp; AM Barrett; Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-1
1a. Impact: H-2; M-2; L-0; I-0; 1b. Performance Gap: H-1; M-2; L-0; I-0
Rationale: One unpublished source cited with few patient encounters included, resource use not noted. However, data drawn from a national outcomes database which, if systematically biased, may represent better outcomes than average. Failure of over 30 percent of patients for whom swallowing was assessed to improve, suggests better use of the measure may identify deficient processes and healthcare disparities which can be addressed.
**In 2011, 7240 (47.(%) on individuals receiving SLP services were treated for a swallowing disorder. This constitutes a large percentage of patients currently receiving services. Of those treated, > 33% failed to progress.There is disparity on those who failed to improve based on insurance, baseline status, diagnosis, and age.
1c. Evidence (*based on decision logic*): Y-3; N-1   IF a **Health Outcome**, rationale supports: Y-2; N-2; NA-0
Quantity: H-1; M-0; L-0; I-1; Quality: H-1; M-0; L-0; I-1; Consistency: H-1; M-0; L-0; I-1
Rationale: Evidence not presented--submitter notes "N/A" as above. Submitter states that supporting evidence linking clinical care with swallowing improvement is of "High" quality without presenting specifics.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-4; N-0
2a. Reliability: H-1; M-3; L-0; I-0; 2b. Validity: H-1; M-3; L-0; I-0
Rationale: Reliability data presented was collected during measure development based on clinical vignettes. Initial evaluation of all clinicians also includes demonstration of > 80 percent reliabilit on written vignettes. Validity was also apparently evaluated at measure development stages, convergently established by comparison with expert clinician panel ratings. Convergent validit with patient satisfaction and self-perceived improvement patient ratings was also somewhat supportive of measure validity.
**Numerator, denominator, exclusions, and risk adjustment clear. Time window ambiguous. Not clear how sampling # was derived. Reliability testing conducted with hypothetical vignettes comparing clinicians to experts. Reliability not tested in a real world setting.   Validity testing - specifications clear; correlation coefficients btw treated clients' satisfaction ratings and clinicians' scores. Large sample. This may help establish MCID of measure. Is this really validity? Maybe more of a responsiveness test. I think this may not be an appropriate validity comparison.

3. Usability: H-2; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: No information presented relevant to quality improvement.
**Measure was developed with input from appropriate stakeholders. Currently used in reporting; recommended by CMS.   However, I don't see a demonstration that the results are currently found to be meaningful, understandable...

4. Feasibility: H-2; M-2; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: Required data elements are available for patients treated by trained SLPs.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-3; N-1
Rationale: Measure implementation so far samples only a small number of stroke patients, and although there is demonstration of reliability

0443 Functional Communicaton Measure: Swallowing

and validity, the rationale for the measure is incomplete. The measure's logic may assume that adherence to a systematic, professional standard links this assessment with excellent execution of specific procedures, better outcomes and methods addressing healthcare disparities, however those steps have not been executed, nor is a plan to execute them articulated. Since only a small fraction of stroke survivors have been assessed with this measure since its inception, the question arises of how the individual and society will see better outcomes of dysphagia care as a result of its maintenance.

Additional Comments/Questions:

Workgroup Call Summary

See notes under measure #0442.

0444 Functional Communication Measure: Spoken Language Expression
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: This measure describes the change in functional communication status subsequent to speech-language pathology treatment related to spoken language expression.
Numerator Statement: Number of stroke patients who make progress as defined by an increase of one or more levels on the Spoken Language Expression Functional Communication Measure (FCM).
Denominator Statement: Number of stroke patients scored on the Spoken Language Expression Functional Communication Measure (FCM).
Exclusions: Patients discharged from speech-language pathology services after only one treatment session. Patients who are not candidates for memory treatment as demonstrated by the highest level of functioning on admission (Level 7 on the Spoken Language Expression Functional Communication Measure).
Adjustment/Stratification: Stratification by risk category/subgroup   N/A.
Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System
Type of Measure: Outcome
Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records
Measure Steward: American Speech-Language-Hearing Association
Workgroup Evaluation Results
The following evaluation ratings and comments are from the Committee Reviewers: Mary Van de Kamp; Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*
Importance to Measure and Report (*based on decision logic*): Y-2; N-1   1a. Impact: H-3; M-0; L-0; I-0; 1b. Performance Gap: H-1; M-1; L-0; I-1   Rationale: In 2011, 6544 (43.3%) of patients receiving SLP services were treated for a spoken language expression disorder and nearly 30% failed to progress. No data on performance gap and data do not indicate disparities in outcome.   1c. Evidence (*based on decision logic*): Y-3; N-0   IF a **Health Outcome**, rationale supports: Y-3; N-0; NA-0   Quantity: H-1; M-0; L-0; I-0; Quality: H-1; M-0; L-0; I-0; Consistency: H-1; M-0; L-0; I-0   Rationale:
2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-0   2a. Reliability: H-2; M-1; L-0; I-0; 2b. Validity: H-1; M-2; L-0; I-0   Rationale: Numerator, denominator, exclusions, target population, risk adjustment model/variables clear. Time window ambiguous. Reliability testing done by comparison of clinicians' ratings of hypothetical patient vignettes to expert ratings. Good to excellent reliability but reliability not tested in a real world environment.   Validity testing reportedly done by comparisons btw clinicians' ratings on measure to satisfaction ratings of treated patients. Developers report that this analysis "found some evidence of convergent, discriminant, and construct validity." The consumer questions are related to patients' satisfaction with service not improvement on the measure. Although correlations btw satisfaction scores and FCM ratings are +, they are not significant. Consumer satisfaction is likely to be influenced by a host of issues other than clinical improvement (as identified in the consumer survey item).
3. Usability: H-2; M-1; L-0; I-0   (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)   Rationale: Measure was developed with input from appropriate stakeholders. Currently used in reporting; recommended by CMS.   However, I don't see a demonstration that the results are currently found to be meaningful, understandable...
4. Feasibility: H-2; M-1; L-0; I-0   (*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)   Rationale: Data elements would be available for clients treated by clinicians who have been trained to use measure.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-3; N-0   Rationale: Additional Comments/Questions:
Workgroup Call Summary
See notes under measure #0442.

0445 Functional Communication Measure: Spoken Language Comprehension

Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure describes the change in functional communication status subsequent to speech-language pathology treatment related to spoken language comprehension.

Numerator Statement: Number of stroke patients who make progress as defined by an increase of one or more levels on the Spoken Language Comprehension Functional Communication Measure (FCM).

Denominator Statement: Number of stroke patients scored on the Spoken Language Comprehension Functional Communication Measure (FCM).

Exclusions: Patients discharged from speech-language pathology services after only one treatment session. Patients who are not candidates for memory treatment as demonstrated by the highest level of functioning on admission (Level 7 on the Spoken Langauge Comprehension Functional Communication Measure).

Adjustment/Stratification: Stratification by risk category/subgroup N/A.

Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System

Type of Measure: Outcome

Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records

Measure Steward: American Speech-Language-Hearing Association

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Mary Van de Kamp; Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-3; N-0
1a. Impact: H-3; M-0; L-0; I-0; 1b. Performance Gap: H-2; M-1; L-0; I-0
Rationale: In 2011, 5592 (37%) of patients receiving SLP services were treated for a spoken language comprehension disorder and 28.4% failed to progress. Disparities on outcome exist based on insurance, baseline data, diagnosis, and age.
1c. Evidence (*based on decision logic*): Y-3; N-0   IF a **Health Outcome**, rationale supports: Y-2; N-1; NA-0
Quantity: H-1; M-0; L-0; I-0; Quality: H-1; M-0; L-0; I-0; Consistency: H-1; M-0; L-0; I-0
Rationale:

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-0
2a. Reliability: H-3; M-0; L-0; I-0; 2b. Validity: H-1; M-2; L-0; I-0
Rationale: Numerator, denominator, exclusions, target population, risk adjustment model/variables clear. Time window ambiguous. Reliability testing done by comparison of clinicians' ratings of hypothetical patient vignettes to expert ratings. Good to excellent reliability but reliability not tested in a real world environment. Validity testing reportedly done by comparisons btw clinicians' ratings on measure to satisfaction ratings of treated patients. Developers report that this analysis "found some evidence of convergent, discriminant, and construct validity." The consumer questions are related to patients' satisfaction with service not improvement on the measure. Although correlations btw satisfaction scores and FCM ratings are +, they are not significant. Consumer satisfaction is likely to be influenced by a host of issues other than clinical improvement (as identified in the consumer survey item).

3. Usability: H-2; M-1; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: Measure was developed with input from appropriate stakeholders. Currently used in reporting; recommended by CMS.   However, I don't see a demonstration that the results are currently found to be meaningful, understandable...

4. Feasibility: H-2; M-1; L-0; I-0
(*4a. Clinical data generated during care process; 4b. Electronic data; 4c. Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)
Rationale: Data are generated by clinicians who are trained and contribute to NOMS.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-3; N-0
Rationale:
Additional Comments/Questions:

Workgroup Call Summary

See notes under measure #0442.

0446 Functional Communicaton Measure: Reading
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: This measure describes the change in functional communication status subsequent to speech-language pathology treatment of patients with reading disorders.
Numerator Statement: Number of stroke patients who make progress as defined by an increase of one or more levels on the Reading Functional Communication Measure (FCM).
Denominator Statement: Number of stroke patients scored on the Reading Functional Communication Measure (FCM).
Exclusions: Patients discharged from speech-language pathology services after only one treatment session. Patients who are not candidates for memory treatment as demonstrated by the highest level of functioning on admission (Level 7 on the Reading Functional Communication Measure).
Adjustment/Stratification: Stratification by risk category/subgroup   N/A
Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System
Type of Measure: Outcome
Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records
Measure Steward: American Speech-Language-Hearing Association
Workgroup Evaluation Results
The following evaluation ratings and comments are from the Committee Reviewers: Mary Van de Kamp; Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*
Importance to Measure and Report (*based on decision logic*): Y-1; N-1   **1a. Impact:** H-2; M-0; L-1; I-0; **1b. Performance Gap:** H-0; M-2; L-0; I-0   <u>Rationale:</u> In 2011, 6544 (16.5%) of patients receiving SLP services were treated for a spoken language expression disorder and nearly 25.4% failed to progress. Disparities on outcome exist based on race, baseline data, and diagnosis. The developer envisions that a benefit of this measure is that "increasing the proportion of patients who make progress on this measure will stimulate clinicians to think about ways to think about ways to improve care to increase this proportion." Couldn't the same be said about most any measure that has a low percent of success. Only public reporting and improvement plans are likely to achieve this.   **1c. Evidence** (*based on decision logic*): Y-3; N-0   IF a **Health Outcome**, rationale supports: Y-3; N-0; NA-0   **Quantity:** H-1; M-0; L-0; I-0; **Quality:** H-1; M-0; L-0; I-0; **Consistency:** H-1; M-0; L-0; I-0   <u>Rationale:</u>
2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-0   **2a. Reliability:** H-1; M-2; L-0; I-0; **2b. Validity:** H-1; M-2; L-0; I-0   <u>Rationale:</u> Numerator, denominator, exclusions, target population, risk adjustment model/variables clear. Time window ambiguous. Reliability testing done by comparison of clinicians' ratings of hypothetical patient vignettes to expert ratings. Good to excellent reliability but reliability not tested in a real world environment.   Validity testing reportedly done by comparisons btw clinicians' ratings on measure to satisfaction ratings of treated patients. Developers report that this analysis "found some evidence of convergent, discriminant, and construct validity." The consumer questions are related to patients' satisfaction with service not improvement on the measure. Although correlations btw satisfaction scores and FCM ratings are +, they are not significant. Consumer satisfaction is likely to be influenced by a host of issues other than clinical improvement (as identified in the consumer survey item).
3. Usability: H-2; M-1; L-0; I-0   (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)   <u>Rationale:</u> Measure was developed with input from appropriate stakeholders. Currently used in reporting; recommended by CMS.   However, I don't see a demonstration that the results are currently found to be meaningful, understandable...
4. Feasibility: H-2; M-1; L-0; I-0   (*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)   <u>Rationale:</u> Data elements would be available for clients treated by clinicians who have been trained to use measure and contribute to NOMS.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-3; N-0   <u>Rationale:</u>   <u>Additional Comments/Questions</u>:
Workgroup Call Summary
See notes under measure #0442.

0447 Functional Communication Measure: Motor Speech
Status: Maintenance, Original Endorsement: Jul 31, 2008

Description: This measure describes the change in functional communication status subsequent to speech-language pathology treatment of patients who exhibit deficits in speech-production.

Numerator Statement: Number of stroke patients who make progress as defined by an increase of one or more levels on the Motor Speech Functional Communication Measure (FCM).

Denominator Statement: Number of stroke patients scored on the the Motor Speech Functional Communication Measure (FCM).

Exclusions: Patients discharged from speech-language pathology services after only one treatment session. Patients who are not candidates for memory treatment as demonstrated by the highest level of functioning on admission (Level 7 on the Motor Speech Functional Communication Measure).

Adjustment/Stratification: Stratification by risk category/subgroup   N/A

Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System

Type of Measure: Outcome

Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records

Measure Steward: American Speech-Language-Hearing Association

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-1; N-0

1a. Impact: H-1; M-0; L-0; I-0; **1b. Performance Gap:** H-0; M-1; L-0; I-0

Rationale: In 2011, 4141 (27%) of patients receiving SLP services were treated for a spoken language expression disorder and nearly 24.1% failed to progress. Disparities on outcome exist based on gender, baseline data, and diagnosis.

1c. Evidence (*based on decision logic*): Y-2; N-0   IF a **Health Outcome**, rationale supports: Y-2; N-0; NA-0

Quantity: H-0; M-0; L-0; I-0; **Quality:** H-0; M-0; L-0; I-0; **Consistency:** H-0; M-0; L-0; I-0

Rationale:

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-2; N-0

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

Rationale: Numerator, denominator, exclusions, target population, risk adjustment model/variables clear. Time window ambiguous. Reliability testing done by comparison of clinicians' ratings of hypothetical patient vignettes to expert ratings. Good to excellent reliability but reliability not tested in a real world environment.   Validity testing reportedly done by comparisons btw clinicians' ratings on measure to satisfaction ratings of treated patients. Developers report that this analysis "found some evidence of convergent, discriminant, and construct validity." The consumer questions are related to patients' satisfaction with service not improvement on the measure. Although correlations btw satisfaction scores and FCM ratings are +, they are not significant. Consumer satisfaction is likely to be influenced by a host of issues other than clinical improvement (as identified in the consumer survey item).

3. Usability: H-1; M-1; L-0; I-0

(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)

Rationale: Measure was developed with input from appropriate stakeholders. Currently used in reporting; recommended by CMS.   However, I don't see a demonstration that the results are currently found to be meaningful, understandable...

4. Feasibility: H-1; M-1; L-0; I-0

(*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)

Rationale: Data elements would be available for clients treated by clinicians who have been trained to use measure and contribute to NOMS.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-2; N-0

Rationale:

Additional Comments/Questions:

Workgroup Call Summary

See notes under measure #0442.

0448 Functional Communication Measure: Memory
Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: This measure describes the change in functional communication status subsequent to speech-language pathology treatment of patients with memory deficits.
Numerator Statement: Number of stroke patients who make progress as defined by an increase of one or more levels on the scale the Memory Functional Communication Measure (FCM).
Denominator Statement: Number of stroke patients scored on the Memory Functional Communication Measure (FCM).
Exclusions: Patients discharged from speech-language pathology services after only one treatment session. Patients who are not a candidate for memory treaments as demonstrated by the highest level of functioning at admission (Level 7 on the Memory Functional Communication Measure).
Adjustment/Stratification: Stratification by risk category/subgroup   N/A
Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System
Type of Measure: Outcome
Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records
Measure Steward: American Speech-Language-Hearing Association

Workgroup Evaluation Results
The following evaluation ratings and comments are from the Committee Reviewers: Mary Van de Kamp; Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*
Importance to Measure and Report (*based on decision logic*): Y-1; N-0   **1a. Impact:** H-2; M-0; L-0; I-0; **1b. Performance Gap:** H-1; M-0; L-0; I-0   <u>Rationale:</u> 1a3 Data available on over 15,000 episodes of care, 32% treated for memory d/o (Good) 1b2 30.1% failure to make progress, random clinician/facility data provided (OK)   **In 2011, 4821 (31.9%) of patients receiving SLP services were treated for a spoken language comprehension disorder and 30.1% failed to progress. Disparities on outcome exist based on gender, race, insurance, baseline data, diagnosis, and age.   **1c. Evidence** (*based on decision logic*): Y-3; N-0   IF a **Health Outcome**, rationale supports: Y-3; N-0; NA-0   **Quantity:** H-1; M-0; L-0; I-0; **Quality:** H-1; M-0; L-0; I-0; **Consistency:** H-1; M-0; L-0; I-0   <u>Rationale:</u> N/A
2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-0   **2a. Reliability:** H-1; M-2; L-0; I-0; **2b. Validity:** H-1; M-2; L-0; I-0   <u>Rationale:</u> 2a1 ? regarding level subjectivity 2b ? expert gold standard concerns   **Numerator, denominator, exclusions, target population, risk adjustment model/variables clear. Time window ambiguous. Reliability testing done by comparison of clinicians' ratings of hypothetical patient vignettes to expert ratings. Good to excellent reliability but reliability not tested in a real world environment.   Validity testing reportedly done by comparisons btw clinicians' ratings on measure to satisfaction ratings of treated patients. Developers report that this analysis "found some evidence of convergent, discriminant, and construct validity." The consumer questions are related to patients' satisfaction with service not improvement on the measure. Although correlations btw satisfaction scores and FCM ratings are +, they are not significant. Consumer satisfaction is likely to be influenced by a host of issues other than clinical improvement (as identified in the consumer survey item).
3. Usability: H-2; M-1; L-0; I-0   (*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)   <u>Rationale:</u> Measure was developed with input from appropriate stakeholders. Currently used in reporting; recommended by CMS.   However, I don't see a demonstration that the results are currently found to be meaningful, understandable...
4. Feasibility: H-3; M-0; L-0; I-0   (*4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented*)   <u>Rationale:</u> 4b. Yes, but combo, not necessarily one place4 4c. There are automated mechanisms   **Data elements would be available for clients treated by clinicians who have been trained to use measure and contribute to NOMS.
Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-3; N-0   <u>Rationale:</u> Measure maintains high or moderate impact ratings for all criteria   <u>Additional Comments/Questions</u>:
Workgroup Call Summary
  See notes under measure #0442.

0449 Functional Communicaton Measure: Attention

Status: Maintenance, Original Endorsement: Jul 31, 2008
Description: This measure describes the change in functional communication status subsequent to speech-language pathology treatment of patients who have attention deficits.
Numerator Statement: Number of stroke patients who make progress as defined by an increase of one or more levels on the Attention Functional Communication Measure (FCM).
Denominator Statement: Number of stroke patients scored on the Attention Functional Communication Measure (FCM).
Exclusions: Patients discharged from speech-language pathology services after only one treatment session. Patients who are not candidates for memory treatment as demonstrated by the highest level of functioning on admission (Level 7 on the Attention Functional Communication Measure).
Adjustment/Stratification: Stratification by risk category/subgroup N/A
Level of Analysis: Clinician : Group/Practice, Facility, Integrated Delivery System
Type of Measure: Outcome
Data Source: Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Registry, Paper Medical Records
Measure Steward: American Speech-Language-Hearing Association

Workgroup Evaluation Results

The following evaluation ratings and comments are from the Committee Reviewers: AM Barrett; Mary Van De Kamp; Jordan Eisenstock; Jane Sullivan   *(comments separated by asterisks)*

Importance to Measure and Report (*based on decision logic*): Y-2; N-1
1a. Impact: H-1; M-1; L-1; I-0; 1b. Performance Gap: H-0; M-3; L-0; I-0
Rationale: Attention disorders are a significant cause of morbidity after stroke, but unfortunately the application does not submit information supporting this impact. Although 18.5% of the patient encounters submitted to ASHA's national outcomes database reported attention disorders, submittors did not link this to morbid outcomes or cost. Performance gap data submitted reflects the variation in recovery of aphasia over treatment periods by SLPs but is not linked directly or indirectly to quality of care in the information submitted.
**In 2011, 2796 (18.5%) of patients receiving SLP services were treated for an attention disorder and 30% failed to progress. Disparities on outcome exist based on gender, race, insurance, baseline data, diagnosis, and age.
1c. Evidence (*based on decision logic*): Y-2; N-1   IF a **Health Outcome**, rationale supports: Y-2; N-1; NA-0
Quantity: H-0; M-0; L-0; I-1; Quality: H-0; M-0; L-0; I-1; Consistency: H-0; M-0; L-0; I-1
Rationale: Although information exists assessing benefits and harms to subjects of health processes coinciding with changes in measures assessing attention (e.g. exposure to medications adversely affecting attention), this evidence is not presented by the submitters. Rather, the submitters present evidence from their database that individual, randomly selected clinicians and individual, randomly selected facilities differ in the proportion of patients which improve, without specifically linking this to health outcomes, benefits, or reduction of harma.

2. Scientific Acceptability of Measure Properties (*based on decision logic*): Y-3; N-0
2a. Reliability: H-0; M-3; L-0; I-0, 2b. Validity: H-0; M-3; L-0; I-0
Rationale: Unpublished data is submitted which may have been collected at the time of measure development. This data reflects a high degree of inter-rater reliability; test-retest reliability is not assessed. It also reflects satisfactory convergent validity of ratings on the measure with expert panel ratings. Face validity was assessed at the time of measure development by expert review (150 SLPs). Formal external validity assessment is presented in unpublished data comparing change with patient-reported satisfaction with care and health outcomes; however, the construct validity of the patient-reported outcome instrument is itself unproven.
**Numerator, denominator, exclusions, target population, risk adjustment model/variables clear. Time window ambiguous. Reliability testing done by comparison of clinicians' ratings of hypothetical patient vignettes to expert ratings. Good to excellent reliability but reliability not tested in a real world environment. Validity testing reportedly done by comparisons btw clinicians' ratings on measure to satisfaction ratings of treated patients. Developers report that this analysis "found some evidence of convergent, discriminant, and construct validity." The consumer questions are related to patients' satisfaction with service not improvement on the measure. Although correlations btw satisfaction scores and FCM ratings are +, they are not significant. Consumer satisfaction is likely to be influenced by a host of issues other than clinical improvement (as identified in the consumer survey item).

3. Usability: H-1; M-2; L-0; I-0
(*Meaningful, understandable, and useful to the intended audiences for 3a. Public Reporting and 3b. Quality Improvement*)
Rationale: Implied in the presentation is that the founding and existence of the database, and reason for evaluating facility-level and individual practitioner-level differences in outcomes on this measure, is to examine differences relevant to quality of care. However, specific examples of how variation from expected valued might be leveraged to identify risk conditions, problems with care processes, or opportunities for healthcare structure improvement, were not discussed.
**Measure was developed with input from appropriate stakeholders. Currently used in reporting; recommended by CMS.   However, I don't see a demonstration that the results are currently found to be meaningful, understandable...

0449 Functional Communicaton Measure: Attention

4. Feasibility: H-1; M-2; L-0; I-0

(4a. Clinical data generated during care process; 4b. Electronic data; 4c.Susceptibility to inaccuracies/ unintended consequences identified 4d. Data collection strategy can be implemented)

<u>Rationale</u>: Positive rating bias (bias toward higher scores) was identified in the reliability assessment. However, relationship of this bias to rater expectancies/bias or other non-random factors is not reported.

**Data elements would be available for clients treated by clinicians who have been trained to use measure and contribute to NOMS.

Preliminary Assessment of Criteria Met/Suitable for Endorsement: Y-2; N-1

<u>Rationale</u>: Unfortunately the assessment is primarily linked to standard of practice for SLPs rather than to specific health outcomes or patient benefits. Research linking use of the assessment or identification of outcome variation with improved health processes or outcomes is needed to support the use of this outcome instrument, since it may not identify as many people with attention problems as other instruments available for cognitive evaluation or delirium assessment (30-40% or more in some studies).

<u>Additional Comments/Questions</u>:

Workgroup Call Summary

See notes under measure #0442.