NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

This form contains the information submitted by measure developers/stewards, organized according to NQF's measure evaluation criteria and process. The evaluation criteria, evaluation guidance documents, and a blank online submission form are available on the submitting standards web page.

NQF #: 1955   **NQF Project:** Neurology Project
(for Endorsement Maintenance Review) **Original Endorsement Date:**   **Most Recent Endorsement Date:**   **Last Updated Date:** Jun 11, 2012
BRIEF MEASURE INFORMATION
De.1 Measure Title: NIH Stroke Scale Recorded
Co.1.1 Measure Steward: American Heart Association/American Stroke Association
De.2 Brief Description of Measure: Percent of patients aged 18 and older with ischemic stroke, or stroke not otherwise specified, with an initial NIH Stroke Scale recorded.
2a1.1 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.
2a1.4 Denominator Statement: Patients with a final clinical diagnosis of ischemic stroke or stroke not otherwise specified.
2a1.8 Denominator 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
1.1 Measure Type: Process **2a1. 25-26 Data Source:** Administrative claims, Electronic Clinical Data : Registry **2a1.33 Level of Analysis:** Facility, Population : National, Population : Regional, Population : State **1.2-1.4 Is this measure paired with another measure?** No **De.3 If included in a composite, please identify the composite measure** (*title and NQF number if endorsed*)**:** Not Applicable.

STAFF NOTES (*issues or questions regarding any criteria*)
Comments on Conditions for Consideration:
Is the measure untested? Yes⦿ No⦿ If untested, explain how it meets criteria for consideration for time-limited endorsement:
1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (*check De.5*)**:** **5. Similar/related endorsed or submitted measures** (*check 5.1*)**:** **Other Criteria:**
Staff Reviewer Name(s):

1. IMPACT, OPPORTUNITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT

Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See guidance on evidence.

Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)

1a. High Impact: H⦿ M⦿ L⦿ I ⦿
(*The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact aspect of healthcare.*)

De.4 Subject/Topic Areas (*Check all the areas that apply*)**:** Neurology : Stroke/Transient Ischemic Attack (TIA)
De.5 Cross Cutting Areas (*Check all the areas that apply*)**:** Access, Disparities, Functional Status, Health and Functional Status : Functional Status

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, A leading cause of morbidity/mortality, Patient/societal consequences of poor quality, Severity of illness

1a.2 If "Other," please describe:

1a.3 Summary of Evidence of High Impact (*Provide epidemiologic or resource use data*)**:**
Stroke is considered the number 4 killer in the United States. It is estimated 7 000 000 Americans >=20 years of age have had a stroke. Each year, >795 000 people experience a new or recurrent stroke. Approximately 610 000 of these are first attacks, and 185 000 are recurrent attacks. On average, every 40 seconds, someone in the United States has a stroke and on average, every 4 minutes, someone dies of a stroke. It is estimated that the direct and indirect cost of stroke in 2008 was $34.3 billion. Given these statistics stroke measures of care are incredibly important.

Demographic variables (age, sex, and insurance status) are not associated with stroke cost. Severe strokes (National Institutes of Health Stroke Scale score >20) cost twice as much as mild strokes, despite similar diagnostic testing. (1-2)

The National Institutes of Health Stroke Scale (NIHSS), which is a validated tool for assessing the initial stroke severity, has been shown to predict mortality in acute ischemic stroke in several prior studies. However, these studies have generally been confined to small numbers of patients from single centers, select patients enrolled in randomized clinical trials, studies outside the United States, or limited to evaluating in-hospital mortality.(3-8)

Recently, a study showed that the NIHSS provides substantial prognostic information regarding 30-day mortality risk in Medicare beneficiaries with acute ischemic stroke. This study concluded that the index of stroke severity is a very strong discriminator of mortality risk, even in the absence of other clinical information, whether used as a continuous or categorical risk determinant.(9)

1a.4 Citations for Evidence of High Impact cited in 1a.3: Citations:

1. Diringer MN, Edwards DF, Mattson DT, Akins PT, Sheedy CW, Hsu CY, Dromerick AW. Predictors of acute hospital costs for treatment of ischemic stroke in an academic center. Stroke. 1999;30:724 –728.

2.Matz R. Cost-effective, risk-free, evidence-based medicine. Arch Intern Med. 2003;163:2795.

3.Johnston KC, Connors AF Jr, Wagner DP, Knaus WA, Wang X, Haley EC Jr. A predictive risk model for outcomes of ischemic stroke.Stroke.2000;31:448–455. Available at: http://stroke.ahajournals.org/content/31/2/448.full.pdf

4. Henon H, Godefroy O, Leys D, Mounier-Vehier F, Lucas C, Rondepierre P,Duhamel A, Pruvo JP. Early predictors of death and disability after acute cerebralischemic event.Stroke. 1995;26:392–398. Available at: http://stroke.ahajournals.org/content/26/3/392.full

5. Nedeltchev K, Renz N, Karameshev A, Haefeli T, Brekenfeld C, Meier N, RemondaL, Schroth G, Arnold M, Mattle HP. Predictors of early mortality after acute ischaemic stroke. Swiss Med Wkly. 2010;140:254–259. Available at: http://www.smw.ch/docs/PdfContent/smw-12919.pdf

6. Chang KC, Tseng MC, Tan TY, Liou CW. Predicting 3-month mortality among patients hospitalized for first-ever acute ischemic stroke. J Formos Med Assoc.2006;105:310–317. Available at: http://www.jfma-online.com/article/S0929-6646(09)60122-4/abstract

7. Weimar C, Konig IR, Kraywinkel K, Ziegler A, Diener HC. Age and National Institutes of Health Stroke Scale score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. Stroke.2004;35:158–162. Available at: http://stroke.ahajournals.org/content/35/1/158.full

8. Smith EE, Shobha N, Dai D, Olson DM, Reeves MJ, Saver JL, Hernandez AF,Peterson ED, Fonarow GC, Schwamm LH. Risk score for in-hospital ischemic stroke mortality derived and validated within the Get With The Guidelines-Stroke Program. Circulation.2010;122:1496–1504. Available at: http://circ.ahajournals.org/content/122/15/1496.full

9. Fonarow G, Saver J, Smith EE, Broderick J, Kleindorfer D, Sacco R, Pan W, Olson D, Hernandez A, Peterson E, Schwamm L. Relationship of National Institute of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries with Acute Ischemic Stroke. J Am Heart Assoc 2012, 1:42-50. Available at: http://jaha.ahajournals.org/content/1/1/42.full

1b. Opportunity for Improvement: H⦿ M⦿ L⦿ I ⦿
(*There is a demonstrated performance gap - variability or overall less than optimal performance*)

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:
The NIHSS is a very strong discriminator of mortality risk, even in the absence of other clinical information, whether used as a continuous or categorical risk determinant. According to one study published categorization of NIHSS into 3 or 4 groups, acute ischemic stroke patients can be readily identified as being at low, medium, or high risk for 30-day mortality, even in the absence of any other demographic or clinical variables. NIHSS provided excellent discrimination of 30-day mortality among multiple clinically relevant subgroups.

1b.2 Summary of Data Demonstrating Performance Gap (*Variation or overall less than optimal performance across providers*)**: [*For Maintenance** – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.*]**
Despite the guideline recommendations and utility for summarizing stroke severity and discriminating outcomes, there are gaps, variations, and disparities in the use of NIHSS. In Get With The Guidelines-Stroke between 37-66% of patients have NIHSS documented.

One recent study examined 33,102 fee-for-service Medicare beneficiaries treated at 404 Get With The Guidelines- Stroke hospitals between April 2003 and December 2006 with NIHSS documented. (1) The 30-day mortality rate by NIHSS as a continuous variable and by risk-tree determined or pre-specified categories were analyzed, with discrimination of risk quantified by the c-statistic. In this cohort, mean age

was 79.0 years and 58% were female. The median NIHSS score was 5 (25th to 75th percentile 2 to 12). There were 4,496 deaths in the first 30 days (13.6%). There was a strong graded relation between increasing NIHSS score and higher 30-day mortality. The 30-day mortality rates for acute ischemic stroke by NIHSS categories were as follows: 0 to 7, 4.2%; 8 to 13, 13.9%; 14 to 21, 31.6%; 22 to 42, 53.5%. A model with NIHSS alone provided excellent discrimination whether included as a continuous variable (c-statistic 0.82 [0.81 to 0.83]), 4 categories (c-statistic 0.80 [0.79 to 0.80]), or 3 categories (c-statistic 0.79 [0.78 to 0.79]).

In assessing the relation of the NIHSS to 30-day mortality after acute ischemic stroke among Get With The Guidelines-Stroke Medicare beneficiaries, this study found that NIHSS is a very strong discriminator of mortality risk. There is a graded near-linear relationship between first recorded NIHSS and higher 30-day mortality risk. This study also demonstrates that with categorization of NIHSS into 3 or 4 groups, acute ischemic stroke patients can be readily identified as being at low, medium, or high risk for 30-day mortality, even in the absence of any other demographic or clinical variables. NIHSS provided excellent discrimination of 30-day mortality among multiple clinically relevant subgroups. NIHSS as a continuous variable or in categories resulted in far better discrimination of 30-day mortality risk then a clinical model not including stroke severity. These findings highlight the importance of a valid specific measure of stroke severity as a determinate of mortality after acute ischemic stroke for Medicare beneficiaries. Further, this study suggests that it may be vital for optimal discrimination to include stroke severity for risk stratification and risk adjustment in acute ischemic stroke.

In this study, a model with NIHSS alone without other variables had a c-statistic of 0.82, a range where clinical risk models are regarded to have genuine clinical utility for individual decision making.(2) This illustrates the very strong relation between stroke severity and 30-day mortality. The mortality risk discrimination capability of NIHSS, either in a continuous or categorical variable, among Medicare beneficiaries with acute ischemic stroke compares very favorably to other mortality prediction models incorporating the NIHSS or other clinical measures of stroke severity (previously reported c-statistics 0.79 to 0.86) and appear superior to models with demographic/clinical variables without severity (previously reported c-statistics 0.69 to 0.75). (3-6) Measures of stroke severity appear to be vital for optimal discrimination of mortality risk.

1b.3 **Citations for Data on Performance Gap: [*For* _Maintenance_ – *Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*]**
Citations:

1. Fonarow G, Saver J, Smith EE, Broderick J, Kleindorfer D, Sacco R, Pan W, Olson D, Hernandez A, Peterson E, Schwamm L. Relationship of National Institute of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries with Acute Ischemic Stroke. J Am Heart Assoc 2012, 1:42-50. Available at: http://jaha.ahajournals.org/content/1/1/42.full

2. Cook NR(2007) Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 115:928–935. Available at: http://circ.ahajournals.org/content/115/7/928.full

3. Weimar C, Konig IR, Kraywinkel K, Ziegler A, Diener HC(2004) Age and National Institutes of Health Stroke Scale score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. Stroke 35:158–162. Available at: http://stroke.ahajournals.org/content/35/1/158.full

4. Smith EE, Shobha N, Dai D, Olson DM, Reeves MJ, Saver JL, Hernandez AF, Peterson ED, Fonarow GC, Schwamm LH(2010) Risk score for in-hospital ischemic stroke mortality derived and validated within the Get With The Guidelines-Stroke Program. Circulation 122:1496–1504.Available at:

http://circ.ahajournals.org/content/122/15/1496.full

5. Fonarow GC, Smith EE, Reeves MJ, Pan W, Olson D, Hernandez AF, Peterson ED, Schwamm LH, Get With The Guidelines Steering Committee and Hospitals(2011) Hospital-level variation in mortality and rehospitalization for medicare beneficiaries with acute ischemic stroke. Stroke 42:159–166. Available at: http://stroke.ahajournals.org/content/42/1/159.full

6. Weimar C, Ziegler A, Konig IR, Diener HC(2002) Predicting functional outcome and survival after acute ischemic stroke. J Neurol 249:888–895. Available at: http://www.springerlink.com/content/lxt6dx2t2017f7t6/

1b.4 Summary of Data on Disparities by Population Group: [_For **Maintenance** –Descriptive statistics for performance results for this measure by population group_**]**
See Attachment. There are two tables.

Table labeled "1b.4 Summary of Data on Disparities by Population Group: Table 1 of 2" was originally published in Relationship of National Institute of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries with Acute Ischemic Stroke(referenced in 1b.5).

Table labeled "1b.4 Summary of Data on Disparities by Population Group: Table: Table 2 of 2" was originally published in Risk score for in-hospital ischemic stroke mortality derived and validated within the Get With The Guidelines–Stroke program(referenced in 1b.5).

Data is not included in this text box becuase it cannot be formatted due to columns/rows.

1b.5 Citations for Data on Disparities Cited in 1b.4: [_For **Maintenance** – Description of the data or sample for measure results reported in 1b.4 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included_**]**
1. Fonarow G, Saver J, Smith EE, Broderick J, Kleindorfer D, Sacco R, Pan W, Olson D, Hernandez A, Peterson E, Schwamm L. Relationship of National Institute of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries with Acute Ischemic Stroke. J Am Heart Assoc 2012, 1:42-50. Available at: http://jaha.ahajournals.org/content/1/1/42.full

2. Smith EE, Shobha N, Dai D, Olson DM, Reeves MJ, Saver JL, Hernandez AF, Peterson ED, Fonarow GC, Schwamm LH(2010) Risk score for in-hospital ischemic stroke mortality derived and validated within the Get With The Guidelines-Stroke Program. Circulation 122:1496–1504.Available at: http://circ.ahajournals.org/content/122/15/1496.full

1c. Evidence (_Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence._)
Is the measure focus a health outcome?   Yes⦿ No⦿   If not a health outcome, rate the body of evidence.

Quantity: H⦿ M⦿ L⦿ I ⦿   Quality: H⦿ M⦿ L⦿ I ⦿   Consistency: H⦿ M⦿ L⦿ I ⦿

Quantity	Quality	Consistency	Does the measure pass subcriterion1c?
M-H	**M-H**	**M-H**	**Yes⦿**
L	**M-H**	**M**	**Yes⦿** IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise **No⦿**
M-H	**L**	**M-H**	**Yes⦿** IF potential benefits to patients clearly outweigh potential harms: otherwise **No⦿**

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

L-M-H	L-M-H	L	No ◉
Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service			**Does the measure pass subcriterion1c?** **Yes**◉ IF rationale supports relationship

1c.1 Structure-Process-Outcome Relationship *(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome)***:**

NIHSS is useful helping to guide acute treatment of stroke. Patients with NIHSS 5-20 (mild to moderate stroke) have the greatest potential for a favorable response to tPA treatment. The chances of a complete recovery among patients with NIHSS >22 improves with treatment but overall success in this group of critically ill patients is low. Because the risk of hemorrhage is considerable among patients with NIHSS >22, the decision to treat with rtPA should be made with caution. Patients with NIHSS <5 may not be candidates for tPA.

There is evidence to support the importance of an initial NIH stroke scale being reported in terms of outcomes. In one recent study, a model with NIHSS alone without other variables had a c-statistic of 0.82, a range where clinical risk models are regarded to have genuine clinical utility for individual decision making.(1) This illustrates the very strong relation between stroke severity and 30-day mortality. The mortality risk discrimination capability of NIHSS, either in a continuous or categorical variable, among Medicare beneficiaries with acute ischemic stroke compares very favorably to other mortality prediction models incorporating the NIHSS or other clinical measures of stroke severity (previously reported c-statistics 0.79 to 0.86) and appear superior to models with demographic/clinical variables without severity (previously reported c-statistics 0.69 to 0.75).(2-5) Measures of stroke severity appear to be vital for optimal discrimination of mortality risk.

There are additional other studies that support the value of the NIHSS however these studies are either smaller, drawn from select patients enrolled in clinical trials, conducted in a small number of centers outside the US, or confined to hospital mortality. Neverethless, they do provide valuable data. In one study of 360 ischemic stroke patients admitted to a single hospital in Taiwan identified admission stroke severity as measured by NIHSS score as the strongest predictor of 3-month mortality, with an odds ratio of 1.17 (95% CI, 1.12–1.22) per point.(6) Another study analyzed 479 patients admitted to a single center in Switzerland, advanced age, and high NIHSS were the only independent predictor of 30-day mortality.(7) A study from 7 centers in Germany found NIHSS obtained within the first 6 hours of admission to be highly predictive of 100-day survival (c-statistic 0.86).(2) A prior study for Get With The Guidelines-Stroke showed that NIHSS was the strongest predictive variable for inhospital mortality and substantially improved the performance of a model based on clinical variables without stroke severity (c-statistic improved from 0.72 to 0.85). (3) However, these studies were either small, drawn from select patients enrolled in clinical trials, conducted in a small number of centers outside the United States, or were confined to inhospital mortality. The findings from the present study, drawn from hundreds of hospitals from all regions of the United States and tens of thousands of patients, substantially extend these prior findings with follow-up to a standard 30- day outcome, and may be viewed as having greater external validity.

Citations:

1.Cook NR(2007) Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 115:928–935. Available at: http://circ.ahajournals.org/content/115/7/928.full

2. Weimar C, Konig IR, Kraywinkel K, Ziegler A, Diener HC(2004) Age and National Institutes of Health Stroke Scale score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. Stroke 35:158–162. Available at: http://stroke.ahajournals.org/content/35/1/158.full

3. Smith EE, Shobha N, Dai D, Olson DM, Reeves MJ, Saver JL, Hernandez AF, Peterson ED, Fonarow GC, Schwamm LH(2010) Risk score for in-hospital ischemic stroke mortality derived and validated within the Get With The Guidelines-Stroke Program. Circulation 122:1496–1504. Available at: Available at: http://circ.ahajournals.org/content/122/15/1496.full

4. Fonarow GC, Smith EE, Reeves MJ, Pan W, Olson D, Hernandez AF, Peterson ED, Schwamm LH, Get With The Guidelines Steering Committee and Hospitals(2011) Hospital-level variation in mortality and rehospitalization for medicare beneficiaries with acute ischemic stroke. Stroke 42:159–166. Available at: Available at: http://stroke.ahajournals.org/content/42/1/159.full

5. Weimar C, Ziegler A, Konig IR, Diener HC(2002) Predicting functional outcome and survival after acute ischemic stroke. J Neurol 249:888–895. Available at: http://www.springerlink.com/content/lxt6dx2t2017f7t6/

6. Chang KC, Tseng MC, Tan TY, Liou CW(2006) Predicting 3-month mortality among patients hospitalized for first-ever acute ischemic stroke. J Formos Med Assoc 105:310–317. Available at: http://www.jfma-online.com/article/S0929-6646(09)60122-4/abstract

7. Nedeltchev K, Renz N, Karameshev A, Haefeli T, Brekenfeld C, Meier N, Remonda L, Schroth G, Arnold M, Mattle HP(2010) Predictors of early mortality after acute ischaemic stroke. Swiss Med Wkly 140:254–259.

Additionally, there is data to show that patients with acute ischemic stroke with NIHSS documented have been shown to be substantially more likely to be treated with TPA than patients without NIHSS documented.

One recent study showed that that many patients not given IV rtPA because of mild or rapidly improving stroke symptoms had poor hospital discharge outcomes[1]. This is the first large, nationwide study to show that patients who do not receive rtPA because of mild or rapidly improving stroke are still at risk for requiring inpatient care or rehabilitation after hospital discharge. Discharge outcomes were significantly worse than for those with TIA. The initial NIHSS score was a strong predictor of outcome, with a graded relation between higher NIHSS score and a lower likelihood of discharge to home or independent ambulation among those with mild or improving ischemic stroke. For the purpose of multivariable modeling, missing data were imputed as follows: age was imputed to the median, categorical variables were imputed to the most common category, and absent past medical history was imputed to "no." Patients with missing data on hospital characteristics were excluded. Because the number with missing NIHSS score was large, yet NIHSS score appeared to be a critical determinant of outcomes, models were run with and without NIHSS score. The median age of the patients with mild/improving stroke was 72 years (interquartile range, 62 to 81 years), and 47.1% were women. The NIHSS score was documented in 18,067 of 29,200 (61.9%) of the patients with mild/improving stroke; the median score was 2 (interquartile range, 1 to 5). Characteristics of mild/ improving stroke patients with versus without a documented NIHSS score are shown in online-only Table III. By comparison, the NIHSS score was documented in 20,095 of 24,292 (82.7%) of the patients given rtPA; the median NIHSS score was 12 (interquartile range, 7 to 18). Patients with higher initial NIHSS scores had worse outcomes (P<0.001, Table 2). Most patients given IV rtPA had NIHSS scores >5 (84.4%), whereas most patients not given IV rtPA because of mild or improving symptoms had NIHSS scores <=5 (80.1%). Patients who were discharged to home had lower initial NIHSS scores and were more likely to be younger, male, and white and to not have diabetes, peripheral vascular disease, or hypertension.

There is data to show that that the NIHSS can be a strong predictor of mortality. In one study researchers analyzed data from 33102 fee-for-service Medicare beneficiaries treated at 404 Get With The Guidelines-Stroke hospitals between April 2003 and December 2006 with NIHSS documented[2]. The 30-day mortality rate by NIHSS as a continuous variable and by risk-tree determined or pre-specified categories were analyzed, with discrimination of risk quantified by the c-statistic. In this cohort, mean age was 79.0 years

and 58% were female. The median NIHSS score was 5 (25th to 75th percentile 2 to 12). There were 4496 deaths in the first 30 days (13.6%). There was a strong graded relation between increasing NIHSS score and higher 30-day mortality. The 30-day mortality rates for acute ischemic stroke by NIHSS categories were as follows: 0 to 7, 4.2%; 8 to 13, 13.9%; 14 to 21, 31.6%; 22 to 42, 53.5%. A model with NIHSS alone provided excellent discrimination whether included as a continuous variable (c-statistic 0.82 [0.81 to 0.83]), 4 categories (c-statistic 0.80 [0.79 to 0.80]), or 3 categories (c-statistic 0.79 [0.78 to 0.79]). Researchers concluded that the NIHSS provides substantial prognostic information regarding 30-day mortality risk in Medicare beneficiaries with acute ischemic stroke. This index of stroke severity is a very strong discriminator of mortality risk, even in the absence of other clinical information, whether used as a continuous or categorical risk determinant.

In another study of GWTG-Stroke data there were 1036 hospitals that contributed 274, 988 ischemic stroke patients to this analyses between October 2001 and December 2007. The sample was randomly divided into a derivation (60%) and validation (40%) sample [3]. Logistic regression was used to determine the independent predictors of mortality and to assign point scores for a prediction model. We also separately derived and validated a model in the 109 187 patients (39.7%) with a National Institutes of Health Stroke Scale (NIHSS) score recorded. Model discrimination was quantified by calculating the C statistic from the validation sample. In-hospital mortality was 5.5% overall and 5.2% in the subset in which NIHSS score was recorded. Characteristics associated with in-hospital mortality were age, arrival mode (eg, via ambulance versus other mode), history of atrial fibrillation, previous stroke, previous myocardial infarction, carotid stenosis, diabetes mellitus, peripheral vascular disease, hypertension, history of dyslipidemia, current smoking, and weekend or night admission. The C statistic was 0.72 in the overall validation sample and 0.85 in the model that included NIHSS score.

A model with NIHSS score alone provided nearly as good discrimination (C statistic 0.83). Plots of observed versus predicted mortality showed excellent model calibration in the validation sample. Researchers concluded that the Get With the Guidelines–Stroke risk model provides clinicians with a well-validated, practical bedside tool for mortality risk stratification. The NIHSS score provides substantial incremental information on a patient's short-term mortality risk and is the strongest predictor of mortality.

Citations:

[1] Smith EE, Fonarow GC, Reeves MJ, Cox M, Olson DM, Hernandez AF, Schwamm LH. Outcomes in mild or rapidly improving stroke not treated with intravenous recombinant tissue-type plasminogen activator: findings from get with the guidelines-stroke. Stroke. 2011 Nov;42(11):3110-5. Epub 2011 Sep 8. Available at: http://stroke.ahajournals.org/content/early/2011/09/08/STROKEAHA.111.613208]

[2] Fonarow, Gregg. Relationship of National Institute of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries with Acute Ischemic Stroke.J Am Heart Assoc 2012;1:42-50. Available at: jaha.ahajournals.org/content/1/1/42.full

[3] [Citation: Smith EE, Nandavar S, Dai D, Olson D, Reeves M, Saver J, Hernandez A, Peterson E, Fonarow G, Schwamm L. A Risk Score for In-Hospital Ischemic Stroke Mortality Derived and Validated within the Get With The Guidelines Stroke Program. Circulation.2010; 122: 1496-1504 Available at: circ.ahajournals.org/content/122/15/1496.full]

For studies that validate the NIHSS instrument, please refer to those studies referenced and summarized in section 1c4. For additional data that shows reliability testing of the NIHSS please refer to Table 3 of "A review of the evidence for the use of telemedicine within stroke systems of care: a scientific statement from the American Heart Association/American Stroke Association."
[Citation: Schwamm LH, Holloway RG, Amarenco P, Audebert HJ, Bakas T, Chumbler NR, Handschu R, Jauch EC, Knight WA 4th, Levine SR, Mayberg M, Meyer BC, Meyers PM, Skalabrin E, Wechsler LR;

American Heart Association Stroke Council; Interdisciplinary Council on Peripheral Vascular Disease. A review of the evidence for the use of telemedicine within stroke systems of care: a scientific statement from the American Heart Association/American Stroke Association. Stroke. 2009 Jul;40(7):2616-34.]

1c.2-3 Type of Evidence *(Check all that apply)***:**
Clinical Practice Guideline

1c.4 Directness of Evidence to the Specified Measure *(State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population)***:**
The NIHSS is useful helping to guide acute treatment of stroke. Patients with NIHSS 5-20 (mild to moderate stroke) have the greatest potential for a favorable response to tPA treatment. The chances of a complete recovery among patients with NIHSS >22 improves with treatment but overall success in this group of critically ill patients is low. Because the risk of hemorrhage is considerable among patients with NIHSS >22, the decision to treat with rtPA should be made with caution. Patients with NIHSS <5 may not be candidates for tPA.

The National Institutes of Health Stroke Scale (NIHSS) is a graded neurological examination that assesses speech, language cognition, inattention, visual field abnormalities, motor and sensory impairments, and ataxia. The scale was developed for use in acute-stroke therapy trials and has since been widely used as a standard part of the assessment in clinical practice. This scale, along with many others, has been evaluated in its clinical usefulness in the assessment of the stroke patient. The ideal stroke scale should be valid, reliable, and easy to administer in multiple settings by a broad range of health care practitioners. The NIHSS has been demonstrated to satisfy all of the criteria for an ideal stroke scale. The NIHSS is not time-consuming to administer, taking <5 minutes to perform. Overall interrater reliability has been shown in multicenter stroke trials and in clinical practice. NIHSS reliability has been extended to non-neurologist physicians, nurses, and stroke coordinators in clinical trials, community-based studies, and large national registries. Factor analysis demonstrated content validity of the NIHSS. Regarding outcomes, the NIHSS has very good sensitivity, specificity, and accuracy in predicting clinical results in-hospital, 30-days, and at 3 months.

Below is a brief synopsis of three of the studies referenced above.
Perhaps the most important study validated the importance of the NIHSS as a means to gauge functional assessment is "Underlying Structure of the National Institutes of Health Stroke Scale Results: a Factor Analysis."In this study research set out to validate as a stroke scale as an outcome measure using data from a clinical trial demonstrating a positive therapeutic effect. Researchers proposed to use data from the National Institute of Neurological Disorders and Stroke (NINDS) tPA Stroke Trial to determine whether the National Institutes of Health Stroke Scale (NIHSS) was valid in patients treated with tissue plasminogen activator (tPA) and to explore the underlying clinimetric structure of the NIHSS. These researchers performed an exploratory factor analysis of NIHSS data from Part 1 (n=291) of the NINDS tPA Stroke Trial to derive a hypothesized underlying factor structure. They then performed a confirmatory factor analysis of this structure using NIHSS data from Part 2 of the same trial (n=333). They tested whether this final factor structure could be found in tPA- and placebo-treated patients serially over time after stroke treatment. Using 3-month outcome data, they tested for an association between the NIHSS and other measures of stroke outcome. The exploratory analysis suggested that there were 2 factors underlying the NIHSS, representing left and right brain function, confirming the content validity of the scale. An alternative structure composed of 4 factors could be derived, with a better goodness of fit: the first 2 factors could represent left brain cortical and motor function, respectively, and the second 2 factors could represent right brain cortical and motor function, respectively. The same factor structures were then found in tPA and placebo patient groups studied serially over time, confirming the exploratory analysis.

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

To assess the predictive validity of the NIHSS using alternative scales, they compared the NIHSS over time with the 3-month outcome using the Barthel Index, Rankin Scale, and Glasgow Outcome Scale. The correlations between the scale and the other clinical outcomes were significant (P<0.001) but modest in magnitude at baseline and 2 hours after stroke. The correlation between the NIHSS at baseline and the measures at 90 days demonstrates predictive validity. The absolute values of the correlation coefficients were greater for the later measurements, suggesting that after 2 hours from stroke, the NIHSS values may have greater predictive validity with respect to the 3-month outcome.

In another study, a single observer scored consecutive admissions to an acute stroke unit on the National Institutes of Health Stroke Scale (NIHSS), the Canadian Neurological Scale, and the Middle Cerebral Artery Neurological Score. Guy´s prognostic score was determined from clinical data. Outcome at 2, 3, 6, and 12 months was categorized as good (alive at home) or poor (alive in care or dead). Predictive accuracy of the variables was compared by receiver operating characteristic curves and stepwise logistic regression. Of the 408 patients studied, 373 had confirmed acute stroke and completed follow-up. The three stroke rating scales each predicted 3-month outcome with an accuracy of .79 or greater. The NIHSS provided the most prognostic information: sensitivity to poor outcome, .71 (95% confidence interval [CI], .64 to .79); specificity, .90 (95% CI, .86 to .94); and overall accuracy, .83 (95% CI, .79 to .87). Logistic regression showed that the NIHSS added significantly to the predictive value of all other scores. No score added useful information to the NIHSS. A cut point of 13 on the NIHSS best predicted 3-month outcome. The researchers concluded that baseline NIHSS best predicts 3-month outcome. The Canadian Neurological Scale and Middle Cerebral Artery Neurological Score also perform well. Baseline assessments in clinical trials only need to include a single stroke rating scale. (2)

In yet another study the reliability of the National Institutes of Health Stroke Scale (NIHSS) for use by trained neurologists in clinical trials of acute stroke has been established in several hospital-based studies. However, it also has the potential for application in community-based settings and to be used by nonneurologists: issues which have not been explored before. Hence, we aimed to determine the reliability of the NIHSS when administered by research nurses within the existing North Eastern Melbourne Stroke Incidence Study. Using the NIHSS, thirty-one consecutively registered stroke patients were assessed by 2 neurologists and 1 of 2 trained research nurses. The interrater reliability of observations was compared using weighted and unweighted kappa statistics and intraclass correlation coefficients (ICC). There was a high level of agreement for total scores between the 2 neurologists (ICC = 0.95) and between each neurologist and research nurse (ICC = 0.92 and 0.96). While there was moderate to excellent agreement among neurologists and research nurse (weighted kappa > 0.4) for the majority of the NIHSS items, there was poor agreement for the component ´limb ataxia´. Overall, agreement between nurse and neurologist for individual items was not significantly different from agreement between neurologists. It appears that in both hospital and community settings, trained research nurses can administer the NIHSS with a reliability similar to stroke-trained neurologists. This ability could be used to advantage in large community-based trials and epidemiological studies. (3)

Citations:
1. Lyden P, Lu M, Jackson C, Marler J, Kothari R, Brott T, Zivin J. Underlying structure of the National Institutes of Health Stroke Scale: results of a factor analysis: NINDS tPA Stroke Trial Investigators. Stroke. 1999;30:2347–2354.

2. Muir KW, Weir CJ, Murray GD, Povey C, Lees KR. Comparison of neurological scales and scoring systems for acute stroke prognosis.Stroke. 1996;27:1817–1820.

3. Goldstein LB, Samsa GP. Reliability of the National Institutes of Health Stroke Scale: extension to non-neurologists in the context of a clinical trial. Stroke. 1997;28:307–310.

4.   Lyden P, Brott T, Tilley B, Welch KM, Mascha EJ, Levine S, Haley EC,Grotta J, Marler J. Improved reliability of the NIH Stroke Scale using video training. Stroke. 1994;25:2220 –2226.

5.   Dewey H, Donnan GA, Freeman EJ, Sharples CM, Macdonell RA, McNeil JJ, Thrift AG. Interrater reliability of the National Institutes of Health Stroke Scale: rating by neurologists and nurses in a community-based stroke incidence study. Cerebrovasc Dis. 1999;9:323–327.

1c.5 Quantity of Studies in the Body of Evidence (*Total number of studies, not articles*)**:** Citations for applicable studies:

1. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med. 1995;333:1581–1587. Available at: http://www.nejm.org/doi/full/10.1056/NEJM199512143332401.

2. Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, Woolson RF, Hansen MD. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: a report of the Trial of ORG 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999;53:126–131. Available at: http://www.neurology.org/content/53/1/126.full

3. Dhamoon MS, Moon YP, Paik MC, Boden-Albala B, Rundek T, Sacco RL, Elkind MS. Long-term functional recovery after first ischemic stroke: the Northern Manhattan Study. Stroke. 2009;40:2805–2811. Available at: http://stroke.ahajournals.org/content/early/2009/06/25/STROKEAHA.109.549576.full.pdf+html

1c.6 Quality of Body of Evidence (*Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events*)**:** The NINDs Trial had two parts. (1) Part 1 (in which 291 patients were enrolled) tested whether t-PA had clinical activity, as indicated by an improvement of 4 points over base-line values in the score of the National Institutes of Health stroke scale (NIHSS) or the resolution of the neurologic deficit within 24 hours of the onset of stroke. Part 2 (in which 333 patients were enrolled) used a global test statistic to assess clinical outcome at three months, according to scores on the Barthel index, modified Rankin scale, Glasgow outcome scale, and NIHSS.(1) In part 2, the long-term clinical benefit of t-PA predicted by the results of part 1 was confirmed (global odds ratio for a favorable outcome, 1.7; 95 percent confidence interval, 1.2 to 2.6). As compared with patients given placebo, patients treated with t-PA were at least 30 percent more likely to have minimal or no disability at three months on the assessment scales. Symptomatic intracerebral hemorrhage within 36 hours after the onset of stroke occurred in 6.4 percent of patients given t-PA but only 0.6 percent of patients given placebo (P < 0.001). Mortality at three months was 17 percent in the t-PA group and 21 percent in the placebo group (P= 0.30).

Another study compared the baseline National Institutes of Health Stroke Scale (NIHSS) score and the Trial of Org 10172 in Acute Stroke Treatment (TOAST) stroke subtype as predictors of outcomes at 7 days and 3 months after ischemic stroke. (2)The baseline NIHSS score strongly predicted outcome, with one additional point on the NIHSS decreasing the likelihood of excellent outcomes at 7 days by 24% and at 3 months by 17%. At 3 months, excellent outcomes were noted in 46% of patients with NIHSS scores of 7 to 10 and in 23% of patients with scores of 11 to 15. After multivariate adjustment, lacunar stroke had an odds ratio of 3.1 (95% CI, 1.5 to 6.4) for an excellent outcome at 3 months.

In yet another study, the population-based Northern Manhattan Study, patients >=40 years of age with incident ischemic stroke were prospectively followed using the Barthel Index at 6 months and annually to 5 years. (3) Baseline stroke severity was categorized as mild (National Institutes of Health Stroke Scale <6),

moderate (6 to 13), and severe (>=14). Follow-up was censored at death, recurrent stroke, or myocardial infarction. Generalized Estimating Equations provided ORs and 95% CIs for predictors of favorable (Barthel Index >=95) versus unfavorable (Barthel Index >=95) functional status after adjusting for demographic and medical risk factors. Of 525 patients, mean age was 68.6+-12.4 years, 45.5% were male, 54.7% Hispanic, 54.7% had Medicaid/no insurance, and 35.1% had moderate stroke. The proportion with Barthel Index >=95 declined over time (OR, 0.91; 95% CI, 0.84 to 0.99). Changes in Barthel Index by insurance status were confirmed by a significant interaction term (ß for Interaction=-0.167, P=0.034); those with Medicaid/no insurance declined (OR, 0.84; P=0.003), whereas those with Medicare/private insurance did not (OR, 0.99; P=0.92).

Citations:

1. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med. 1995;333:1581–1587. Available at: http://www.nejm.org/doi/full/10.1056/NEJM199512143332401.

2. Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, Woolson RF, Hansen MD. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: a report of the Trial of ORG 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999;53:126–131. Available at: http://www.neurology.org/content/53/1/126.full

3. Dhamoon MS, Moon YP, Paik MC, Boden-Albala B, Rundek T, Sacco RL, Elkind MS. Long-term functional recovery after first ischemic stroke: the Northern Manhattan Study. Stroke. 2009;40:2805–2811. Available at: http://stroke.ahajournals.org/content/early/2009/06/25/STROKEAHA.109.549576.full.pdf+html

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):* Assessment of stroke severity is one of the essential steps in determining eligibility for intravenous tPA for acute ischemic stroke which in term have been demonstrated to improve clinical outcomes. The number needed to treat for substantial benefit with IV tPA in acute ischemic stroke ranges from 3.5 to 5.9 depending on onset to treatment time.

Patients with acute ischemic stroke with NIHSS documented have been shown to be substantially more likely to be treated with tPA than patients without NIHSS documented.

NIHSS is the most important determinate of in-hospital mortality, 30-day mortality, and 1 year mortality and stroke functional outcomes at 90 days. It has been shown to be the strongest predictor of outcome and therefore is a fundamental factor in risk adjustment.

1c.8 Net Benefit *(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms)***:**
NIHSS is useful helping to guide acute treatment of stroke. Patients with NIHSS 5-20 (mild to moderate stroke) have the greatest potential for a favorable response to tPA treatment. The chances of a complete recovery among patients with NIHSS >22 improves with treatment but overall success in this group of critically ill patients is low. Because the risk of hemorrhage is considerable among patients with NIHSS >22, the decision to treat with rtPA should be made with caution. Patients with NIHSS <5 may not be candidates for tPA.

Determination and reporting of NIHSS is essential for identification of patients eligibility for certain therapies which in turn improve acute ischemic stroke outcomes, providing prognostic information to clinicians, patients, and family members regarding function outcomes, in-hospital mortality, 30-day mortality and 1-year mortality, and reporting of hospital-level risk standardized outcome measures. NIHSS is also one of the key factors in estimating the likelihood of good outcome after tPA treatment and risk of complications with

tPA.

Assessment of stroke severity is one of the essential steps in determining eligibility for intravenous tPA for acute ischemic stroke which in term have been demonstrated to improve clinical outcomes. The number needed to treat for substantial benefit with IV tPA in acute ischemic stroke ranges from 3.5 to 5.9 depending on onset to treatment time.

Patients with acute ischemic stroke with NIHSS documented have been shown to be substantially more likely to be treated with tPA than patients without NIHSS documented.

NIHSS is the most important determinate of in-hospital mortality, 30-day mortality, and 1 year mortality and stroke functional outcomes at 90 days. It has been shown to be the strongest predictor of outcome and therefore is a fundamental factor in risk adjustment.

Because of the widespread acceptance of the NIHSS, its use in making clinical decisions and in determining eligibility for certain therapies, and the need for quantified measures of initial severity to interpret data about outcomes and other metrics, this is an essential metric for stroke quality of care.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? Yes

1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: American Heart Association/American Stroke Association through its guidelines. The AHA/ASA is the submitter of this measure.

1c.11 System Used for Grading the Body of Evidence: Other

1c.12 If other, identify and describe the grading scale with definitions: Indications are categorized as Class I, II, or III on the basis of a multifactorial assessment of risk and expected efficacy viewed in the context of current knowledge and the relative strength of this knowledge. These classes summarize the recommendations for procedures or treatments as follows:

Class I: Conditions for which there is evidence for and/or general agreement that a given procedure or treatment is beneficial, useful, and effective.

Class II: Conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment.
Class IIa: Weight of evidence/opinion is in favor of usefulness/efficacy.
Class IIb: Usefulness/efficacy is less well established by evidence/opinion.

Class III: Conditions for which there is evidence and/or general agreement that a procedure/treatment is not useful/effective and in some cases may be harmful.

1c.13 Grade Assigned to the Body of Evidence: Class I

1c.14 Summary of Controversy/Contradictory Evidence: The NIH Stroke Scale is not the only scoring system to determine acute stroke prognosis. Other scales/scoring systems that provide prognostic information include Canadian Neurological Scale, Middle Cerebral Artery Neurological Scale, Scandinavian Stroke Scale, European Stroke Scale and Guy's score. Use of the NIHSS requires specialized training.

1c.15 Citations for Evidence other than Guidelines*(Guidelines addressed below)***:**
1. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med. 1995;333:1581–1587. Available at:

http://www.nejm.org/doi/full/10.1056/NEJM199512143332401.

2. Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, Woolson RF, Hansen MD. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: a report of the Trial of ORG 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999;53:126–131. Available at: http://www.neurology.org/content/53/1/126.full

3. Dhamoon MS, Moon YP, Paik MC, Boden-Albala B, Rundek T, Sacco RL, Elkind MS. Long-term functional recovery after first ischemic stroke: the Northern Manhattan Study. Stroke. 2009;40:2805–2811. Available at: http://stroke.ahajournals.org/content/early/2009/06/25/STROKEAHA.109.549576.full.pdf+html

1c.16 Quote verbatim, <u>the specific guideline recommendation</u> *(Including guideline # and/or page #)***:**
The use of a stroke rating scale, preferably the NIHSS, is recommended. Hospitals (ie, administration) must provide the necessary resources to use such a scale.(Class I; Level of Evidence B). Stroke2007, 38:1655-1711; Page 1666.

Guideline strongly recommends that the patient be assessed for stroke severity using the NIHSS at the time of presentation/hospital admission, or at least within the first 24 hours after presentation.Stroke 2005, 36:e100-e143; Page e106.

Strongly recommend that the patient be assessed for stroke severity using the NIHSS at the time of presentation/ hospital admission, or at least within the first 24 hours after presentation.Stroke 2005, 36:e100-e143; Page e113

Strongly recommend that all professionals involved in any aspect of the stroke care be trained and certified to assess stroke severity using the NIHSS.Stroke 2005, 36:e100-e143; Page e113.

Guidelines strongly recommends that patients be reassessed using the NIHSS at the time of acute care discharge.Stroke 2005, 36:e100-e143; Page e113.

1c.17 Clinical Practice Guideline Citation: Adams HP Jr, del Zoppo G, Alberts MJ, Bhatt DL, Brass L, Furlan A, Grubb RL, Higashida RT, Jauch EC, Kidwell C, Lyden PD, Morgenstern LB, Qureshi AI, Rosenwasser RH, Scott PA, Wijdicks EFM. Guidelines for the early management of adults with ischemic stroke: a guideline from the American Heart Association/ American Stroke Association Stroke Council, Clinical Cardiology Council,
Cardiovascular Radiology and Intervention Council, and the Atherosclerotic Peripheral Vascular Disease and Quality of Care Outcomes in Research Interdisciplinary Working Groups. Stroke. 2007; 38: 1655–1711. Available at:
http://stroke.ahajournals.org/content/38/5/1655.full.pdf+html.

Duncan PW, Zorowitz R, Bates B, Choi JY, Glasberg JJ, Graham GD, et al. Management of adult stroke rehabilitation care: a clinical practice guideline. Stroke. 2005;36:e100–143. Availabel at: http://stroke.ahajournals.org/content/36/9/e100.full

1c.18 National Guideline Clearinghouse or other URL: See URLs above in 1c.17 (not fit in this box)

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? Yes

1c.20 If guideline recommendation graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: American Heart Association/American Stroke Association grade the recommendations. The AHA/ASA is submitting the measure.

1c.21 System Used for Grading the Strength of Guideline Recommendation: Other

1c.22 If other, identify and describe the grading scale with definitions: -Level of Evidence A: Data derived from multiple randomized clinical trials or meta-analyses.
-Level of Evidence B: Data derived from a single randomized trial or nonrandomized studies.
-Level of Evidence C: Only consensus opinion of experts, case studies, or standard-of-care.

1c.23 Grade Assigned to the Recommendation: Level of Evidence B

1c.24 Rationale for Using this Guideline Over Others: These guidelines are the most widely recognized professional guidelines in the US for treatment of patients with stroke and TIA.

Based on the NQF descriptions for rating the evidence, what was the <u>developer's assessment</u> of the quantity, quality, and consistency of the body of evidence?
1c.25 Quantity: High   **1c.26** Quality: High**1c.27** Consistency**:** High
1c.28 Attach evidence submission form**:**
1c.29 Attach appendix for supplemental materials**:**

Was the threshold criterion, *Importance to Measure and Report*, met?
(*1a & 1b must be rated moderate or high and 1c yes*)   Yes⊙ No⊙
Provide rationale based on specific subcriteria:

For a new measure if the Committee votes NO, then STOP.
For a measure undergoing endorsement maintenance, if the Committee votes NO because of 1b. (no opportunity for improvement), it may be considered for continued endorsement and all criteria need to be evaluated.

2. RELIABILITY & VALIDITY - SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

Extent to which the measure, <u>as specified</u>, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. **(evaluation criteria)**
Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate field. Supplemental materials may be referenced or attached in item 2.1. See guidance on measure testing.

S.1 Measure Web Page *(In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained)*. Do you have a web page where current detailed specifications for <u>this</u> measure can be obtained? No

S.2 If yes, provide web page URL:

2a. RELIABILITY. Precise Specifications and Reliability Testing: H⊙ M⊙ L⊙ I ⊙

2a1. Precise Measure Specifications. *(The measure specifications precise and unambiguous.)*

2a1.1 Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome)***:**
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.

2a1.2 Numerator Time Window *(The time period in which the target process, condition, event, or outcome is eligible for inclusion)***:**

2a1.3 Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
• Initial NIH Stroke Scale: Yes (Whole integer numerical value between 0-42)
• NIHSS Total Score: is NOT blank

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
Patients with a final clinical diagnosis of ischemic stroke or stroke not otherwise specified.

2a1.5 Target Population Category *(Check all the populations for which the measure is specified and tested if any)*: Adult/Elderly Care

2a1.6 Denominator Time Window *(The time period in which cases are eligible for inclusion)*:
During the period of hospitalization.

2a1.7 Denominator Details (*All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
An ICD-9-CM/ICD-10 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 strokeICD-10 :I6322, I6529, I63139,I63239, I63019,I63119,I63219,I6359,I6359,I6320,I6609,I6619,I6629, I6330,I6340,I6350, I678.

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
• 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

2a1.9 Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
• Age<18
• Patient location when stroke symptoms discovered
• Stroke symptoms resolved at time of presentation
• Patient underwent elective carotid intervention

2a1.10 Stratification Details/Variables *(All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses)*:
Not Applicable.

2a1.11 Risk Adjustment Type *(Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13)*: No risk adjustment or risk stratification   **2a1.12 If "Other," please describe:**

2a1.13 Statistical Risk Model and Variables *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.)*:
Not Applicable.

2a1.14-16 Detailed Risk Model Available at Web page URL (or attachment). Include coefficients, equations, codes with descriptors, definitions, and/or specific data collection items/responses. Attach

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

documents only if they are not available on a webpage and keep attached file to 5 MB or less. NQF strongly prefers you make documents available at a Web page URL. Please supply login/password if needed:

2a1.17-18. Type of Score: Rate/proportion

2a1.19 Interpretation of Score (*Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score*):
Better quality = Higher score

2a1.20 Calculation Algorithm/Measure Logic(*Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.*):
Rate is determined by calculating those eligible patients meeting the numerator specification divided by those meeting the denominator specification.

For this measure:
1) Check to see if there is an ICD-9/ICD-10 principal diagnosis of ischemic stroke or stroke not otherwise specified, exclude those patients not on list
2) Check to see if patient had an inpatient stroke, exclude those patients with inpatient stroke
3) Check to see if patient is 18 years or older; exclude those patients 18 or younger
4) Check to see if patient had an elective carotid intervention, exclude those patients admitted for sole purpose of performance of elective carotid intervention
5) Check to see if patient's stroke symptoms had resolved at time of presentation; if symptoms resolved at time of presentation, exclude patient
6) Check to see if an initial NIH Stroke Scale and NIHSS Total Score were recorded. If NIHSS was not recorded, keep patient in measure. If an NIH Stroke Scale was performed and the Total Score was 0-42 then include patient; exclude those patients with blank Total NIHSS Score.

For detailed measure algorithm see attached.

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:
Attachment
Specifications_2a1.21_NQF_NIHSS_Recorded_Flow_Chart_Version_2-634716649027633418.pdf

---

2a1.24 Sampling (Survey) Methodology. If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):
Not Applicable.

2a1.25 Data Source (*Check all the sources for which the measure is specified and tested*). If other, please describe:
 Administrative claims, Electronic Clinical Data : Registry

2a1.26 Data Source/Data Collection Instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*): Get With The Guidelines-Stroke Registry

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment: URL

---

http://www.heart.org/idc/groups/heart-public/@wcm/@private/@hcm/@gwtg/documents/downloadable/ucm_432072.pdf

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:
Attachment
Specifications_2a1.30_Importance_1b.4.Feasibility_4d.1_NIHSS_Recorded.pdf

2a1.33 Level of Analysis (*Check the levels of analysis for which the measure is specified and tested*)**:**
Facility, Population : National, Population : Regional, Population : State

2a1.34-35 Care Setting (*Check all the settings for which the measure is specified and tested*)**:**
Hospital/Acute Care Facility

2a2. Reliability Testing. (*Reliability testing was conducted with appropriate method, scope, and adequate demonstration of reliability*.)

2a2.1 Data/Sample (*Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*)**:**
As of April, 2012, GWTG-Stroke had 1,493 hospitals submitting data, with 2,030,279 records submitted. A stroke module data quality audit was conducted in 2010 that consisted of 438 medical records randomly selected by SAS from 147 GWTG-Stroke hospitals, with records from each hospital containing at least one tPA patient. The data submitted to the GWTG-Stroke program were compared against the medical record by a trained coder at the independent statistical coordinating center. No significant differences among participating hospitals were found in overall Inter-rater reliability by bed size, ischemic stroke volume, primary stroke center certification, or Coverdell Registry participation.

2a2.2 Analytic Method (*Describe method of reliability testing & rationale*)**:**
The NIHSS has been validated by Goldstein, et al (see below citation). In the GWTG-Stroke data, this variable has a very low missing rate (< 3%.). Among the variables used to calculate this measure, accuracy and inter-rater reliability was high with all evaluated variables above 0.84 and most above 0.92. Variables not included in the audit, but included in the measure, included NIHSS score, patient location where stroke symptoms discovered, clinical trial status, and admission for elective carotid intervention. In an effort to minimize site burden (note all participating sites volunteer to participate in GWTG-Stroke) for this off-site audit, blinded copies of all medical records were not requested. Therefore, a 1:1 audit of every variable was not possible. Additionally, sites were not informed a priori of fields for which reliability would be assessed. Therefore, requests for records for the audit were made by identifying the type of record, as opposed to the field itself.

[Goldstein LB, Bertels C, Davis JN. Interrater reliability of the NIH Stroke Scale. Arch Neurol. 1989; 46:660-662.]

2a2.3 Testing Results (*Reliability statistics, assessment of adequacy in the context of norms for the test conducted*)**:**
Initial NIHSS performed was found to have an accuracy of 93.6%, with an excellent inter-rater reliability of 0.89 (kappa statistic, 95% confidence interval 0.79-0.91.) The NIHSS has been validated by Goldstein, et al. In the GWTG-Stroke data, this variable has a very low missing rate (< 3%.). Among the variables used to calculate this measure, accuracy and inter-rater reliability was high with all evaluated variables above 0.84 and most above 0.92. Variables not included in the audit, but included in the measure, included NIHSS score, patient location where stroke symptoms discovered, clinical trial status, and admission for elective carotid intervention. In an effort to minimize site burden (note all participating sites volunteer to participate in GWTG-Stroke) for this off-site audit, blinded copies of all medical records were not requested.

Therefore, a 1:1 audit of every variable was not possible. Additionally, sites were not informed a priori of fields for which reliability would be assessed. Therefore, requests for records for the audit were made by identifying the type of record, as opposed to the field itself.

2b. VALIDITY. Validity, Testing, including all Threats to Validity: H⦾ M⦾ L⦾ I ⦾

2b1.1 Describe how the measure specifications (measure focus, target population, and exclusions) **are consistent with the evidence cited in support of the measure focus (**criterion 1c**) and identify any differences from the evidence:**

Percent of patients aged 18 and older with ischemic stroke, or stroke not otherwise specified with an initial NIH Stroke Scale recorded is the focus area of the measure. The National Institutes of Health Stroke Scale (NIHSS), which is a validated tool for assessing the initial stroke severity, has been shown to predict mortality in acute ischemic stroke in several prior studies. Furthermore, this recommendation is consistent with ASA guidelines related to the value in having stroke patient assessed for stroke severity using the NIHSS at the time of presentation/ hospital admission, or at least within the first 24 hours after presentation, and that all professionals involved in any aspect of the stroke care be trained and certified to assess stroke severity using the NIHSS. This measure however has been drafted to include detailed list of exclusions to prevent any unintended consequences associated with the measure.

There are a number of studies that examine how diverse individuals can be trained in the NIHSS. Please refer to those studies referenced in 1c.4.

Moreover, as noted in earlier sections of this form, there are studies that have also shown the benefits of patients with acute ischemic stroke with NIHSS documented have been shown to be substantially more likely to be treated with TPA than patients without NIHSS documented. To see a summary of those studies, please refer to last three studies referenced at the bottom of 1c.1.

There are number of published studies that show that the NIHSS can be performed by non-neurologists, and which show that the training can be undertaken by a broad range of individuals. Below are the citations as well as brief summary of the study.

1.Goldstein L, and Samsa, G. Reliability of the National Institutes of Health Stroke Scale. Stroke. 1997;28:307. Available at: http://stroke.ahajournals.org/content/28/2/307.abstract
Summary: In this study, the initiation of a randomized trial of a new therapy for patients with acute ischemic stroke, 30 physician investigators (30% of whom were not neurologists) and 29 non-physician study coordinators were trained in the use of the NIHSS at an informational and training conference using standardized videotaped patient examinations. A series of 4 patients were rated initially. After 3 months, the same 4 patients were rerated, providing a measure of intraobserver reliability. An additional series of 4 new patients were also rated after 3 months and, with the initial 4 ratings, provided data for assessment of interobserver reliability. Overall, 28% of the raters had previous experience with the NIHSS, and 22% had previously used the videotapes as used in the present trial. The coefficients of determination ($r2$) were each greater than .95 when the means of the two ratings of the same 4 cases were compared between (1) neurologists and other types of physicians, (2) physicians and study coordinators, (3) raters who had prior experience with the NIHSS and those without prior experience, and (4) raters who had used the videotapes in the past and those who had never viewed the tapes. The calculated $r2$s were greater than .98 for the initial rating of the first 4 cases and for the later rating of the 4 new cases. The slopes of the regression lines were all near 1, indicating that the raters were similarly calibrated. The intraclass correlation coefficients were .93 and .95, reflecting high levels of intraobserver and interobserver reliability. This study concluded, these data extend the previously demonstrated reliability of the NIHSS to non-neurologists and show that both a variety of physician investigators and nurse study coordinators can be rapidly trained to reliably apply the scale in the context of an actual clinical trial.

2.Lyden P, Brott T, Tilley B, Welch KM, Mascha EJ, Levine S, Haley EC, Grotta J, Marler J. Improved

reliability of the NIH Stroke Scale using video training. Stroke. 1994;25:2220–2226. Available at:
http://stroke.ahajournals.org/content/25/11/2220.full.pdf
Summary: In this study, the researchers desired to train investigators to use the National Institutes of Health Stroke Scale in a study of acute stroke therapy so that all examiners rated patients comparably. Researchers devised a two-camera videotape method that optimizes the visual presentation of examination findings and measured the effectiveness of the training by asking each investigator to evaluate a set of 11 patients, also on videotape. Researchers found moderate to excellent agreement on most Stroke Scale items (unweighted kappa > 0.60). Two items, facial paresis and ataxia, exhibited poor agreement (unweighted kappa < 0.40) and should be revised in future editions of the scale. Performance improved with video training compared with previous studies. Inclusion of the motor rating of the unaffected limbs in the total score did not affect reliability. This study concluded that video training and certification is a practical and effective method to standardize the use of examination scales. Two cameras must be used during the taping of patients to accurately present the clinical findings. This method is easily adapted to any study in which a large number of investigators will be enrolling patients at multiple clinical centers.

3.Goldstein LB, Bartels C, Davis JN. Interrater reliability of the NIH Stroke Scale. Arch Neurol. 1989;46:660. Available at:
http://archneur.jamanetwork.com/article.aspx?volume=46&issue=6&page=660
Summary: The interobserver reliability of a rating scale employed in several multicenter stroke trials was investigated. Twenty patients who had a stroke were rated with this scale by four clinical stroke fellows. Each patient was independently evaluated by one pair of observers. The degree of interrater agreement for each item on the scale was determined by calculation of the kappa statistic. Interobserver agreement was moderate to substantial for 9 of 13 items. This rating system compares favorably with other scales for which such comparisons can be made. However, the validity of this system must be established.

4.Albanese MA, Clarke WR, Adams HP Jr, Woolson RF. Ensuring reliability of outcome measures on multicenter clinical trials of treatments for acute ischemic stroke: the program developed for the Trial of ORG 10172 in Acute Stroke treatment (TOAST). Stroke. 1994;25:1746. Available at:
http://stroke.ahajournals.org/content/25/9/1746.abstract
Summary: This paper describes an approach to train physicians to use three clinical measures: the National Institutes of Health (NIH) Stroke Scale, a supplemental motor examination, and the Glasgow Outcome Scale. The program included education, certification, remediation when needed, monitoring, and reliability assessment. The goal was to ensure that interrater assessments were as equivalent to one another as possible. Of the first 95 clinicians who began the certification process, 75 passed during the first evaluation. Eighteen of the other physicians were able to complete the process after remediation. The intraclass correlations of both the NIH Stroke Scale and supplemental motor examination exceeded 0.95. The kappa values for the Glasgow Outcome Scale were 0.61 and 0.62 for the first and second ratings of the videotape, respectively. The researchers experience suggested that a program that includes educational and certification processes can be performed as part of the design of a multicenter clinical trial. The method of providing educational and testing videotapes to each site so that physicians can be trained and certified is an effective, inexpensive, and practical approach for enhancing and certifying the expertise of the large number of physicians involved in a multicenter study.

5.Dewey H, Donnan GA, Freeman EJ, Sharples CM, Macdonell RA, McNeil JJ, Thrift AG. Interrater reliability of the National Institutes of Health Stroke Scale: rating by neurologists and nurses in a community-based stroke incidence study. Cerebrovasc Dis. 1999; 9: 323–327. Available at:
http://www.mendeley.com/research/interrater-reliability-national-institutes-health-stroke-scale-rating-neurologists-nurses-communitybased-stroke-incidence-study/
Summary: The reliability of the National Institutes of Health Stroke Scale (NIHSS) for use by trained neurologists in clinical trials of acute stroke has been established in several hospital-based studies. However, it also has the potential for application in community-based settings and to be used by nonneurologists: issues which have not been explored before. Hence, we aimed to determine the reliability

of the NIHSS when administered by research nurses within the existing North Eastern Melbourne Stroke Incidence Study. Using the NIHSS, thirty-one consecutively registered stroke patients were assessed by 2 neurologists and 1 of 2 trained research nurses. The interrater reliability of observations was compared using weighted and unweighted kappa statistics and intraclass correlation coefficients (ICC). There was a high level of agreement for total scores between the 2 neurologists (ICC = 0.95) and between each neurologist and research nurse (ICC = 0.92 and 0.96). While there was moderate to excellent agreement among neurologists and research nurse (weighted kappa > 0.4) for the majority of the NIHSS items, there was poor agreement for the component ´limb ataxia´. Overall, agreement between nurse and neurologist for individual items was not significantly different from agreement between neurologists. It appears that in both hospital and community settings, trained research nurses can administer the NIHSS with a reliability similar to stroke-trained neurologists. This ability could be used to advantage in large community-based trials and epidemiological studies.

6.Richardson J, Murray D, House CK, Lowenkopf T. Successful implementation of the National Institutes of Health Stroke Scale on a stroke/neurovascular unit. J Neurosci Nurs. 2006 Sep;38(4 Suppl):309-15.
Available at:
http://journals.lww.com/jnnonline/Abstract/2006/09000/Successful_Implementation_of_the_National.7.aspx

Summary: The National Institutes of Health Stroke Scale (NIHSS) is accepted as the definitive clinical examination to assess stroke severity. This project examined barriers to implementation and NIHSS use by registered nurses on a stroke/neurovascular Unit. Staff members were surveyed to determine nurse-perceived barriers to the routine use of the NIHSS. Survey results were used to create interventions including staff education, emphasis on NIHSS assessment during interdisciplinary rounds, and use of pocket cards. When the survey was redistributed 9 months later to verify results of the quality improvement initiative and guide further interventions, NIHSS assessment had increased from 12% to 69%. NIHSS scores have been linked to an existing outcomes database to monitor acute stroke treatment and inpatient management outcomes.

7.Goldstein LB, Samsa GP. Reliability of the National Institutes of Health Stroke Scale: extension to non-neurologists in the context of a clinical trial. Stroke.. 1997;28:307–310. Available at:
http://stroke.ahajournals.org/content/28/2/307.long
Summary: In this study the intraclass correlation coefficients were .93 and .95, reflecting high levels of intraobserver and interobserver reliability. These data extend the previously demonstrated reliability of the NIHSS to non-neurologists and show that both a variety of physician investigators and nurse study coordinators can be rapidly trained to reliably apply the scale in the context of an actual clinical trial.

8. Lyden P, Ramani R, Liu L, Emr M, Warren M, Marler J. NIHSS certification is reliable across multiple venues. Stroke. 2009;40:2507–2511. Available at:
http://stroke.ahajournals.org/content/40/7/2507.full?maxtoshow=&hits=10&RESULTFORMAT=&fulltext=nihss&searchid=1&FIRSTINDEX=150&resourcetype=HWFIG
Summary: This study evaluated 8214 health care providers: 33% of all responses came from registered nurses, 23% from emergency department MD/other emergency department/other physicians, and 44% from neurologists. The intraclass correlation coefficient for total score was 0.85 (95% CI, 0.72 to 0.90). Reliability scores were similar among specialists and there were no major differences between nurses and physicians.

9. Kasner SE, Chalela JA, Luciano JM, Cucchiara BL, Raps EC, McGarvey ML,Conroy MB, Localio AR. Reliability and validity of estimating the NIH stroke scale score from medical records. Stroke. 1999;30:1534–1537. Available at: http://stroke.ahajournals.org/content/30/8/1534.abstract
Summary: The NIHSS can be abstracted from medical records with a high degree of reliability and validity. Interrater reliability was excellent, with an intraclass correlation coefficient of 0.82. Scores were well calibrated among the raters.

10.Lyden P, Raman R, Liu L, Grotta J, Broderick J, Olson S, Shaw S, Spilker S, Meyer B, Emr M, Warren M, Marler J. NIHSS training and certification using a new digital video disk is reliable. Stroke. 2005;36:2446–2449. Available at: http://stroke.ahajournals.org/content/36/11/2446
Summary:
After producing a new NIHSS training and demonstration DVD, researchers selected 18 patients representing all possible scores on 15 scale items for a new certification DVD. Patients were divided into 3 certification groups of 6 patients each, balanced for lesion side, distribution of scale item findings, and total score. Researchers sought to measure interrater reliability of the certification DVD using methodology previously published for the original videotapes. Raters were recruited from 3 experienced stroke centers. Each rater watched the new training DVD and then evaluated one of the 3 certification groups. Responses were received from 112 raters: 26.2% of all responses came from stroke nurses, 34.1% from emergency departments/other physicians, and 39.6% from neurologists. One half (50%) of raters were previously NIHSS-certified. Item responses were tabulated, scoring performed as previously published, and agreement measured with unweighted kappa coefficients for individual items and an intraclass correlation coefficient for the overall score. kappa ranged from 0.21+/-0.05 (ataxia) to 0.92+/-0.09 (LOC-C questions). Of 15 items, 2 showed poor, 11 moderate, and 2 excellent agreement based on kappa scores. The intraclass correlation coefficient for total score was 0.94 (95% confidence interval, 0.84 to 1.00). Reliability scores were similar among specialists and centers, and there were no differences between nurses and physicians kappa scores trended higher among raters previously certified. Researchers concluded certification DVDs are reliable for NIHSS certification, and scoring sheets have been posted on a web site for real-time, online certification.

2b2. Validity Testing. (*Validity testing was conducted with appropriate method, scope, and adequate demonstration of validity.*)

2b2.1 Data/Sample (*Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*)**:**
We analyzed data from 33,102 fee-for-service Medicare beneficiaries treated at 404 Get With The Guidelines-Stroke hospitals between April 2003 and December 2006 with NIHSS documented. The 30-day mortality rate by NIHSS as a continuous variable and by risk-tree determined or prespecified categories were analyzed, with discrimination of risk quantified by the c-statistic. In this cohort, mean age was 79.0 years and 58% were female. The median NIHSS score was 5 (25th to 75th percentile 2 to 12). There were 4496 deaths in the first 30 days (13.6%). There was a strong graded relation between increasing NIHSS score and higher 30-day mortality. The 30-day mortality rates for acute ischemic stroke by NIHSS categories were as follows: 0 to 7, 4.2%; 8 to 13, 13.9%; 14 to 21, 31.6%; 22 to 42, 53.5%. In the harvest data, this variable has a very low missing rate (< 3%.) Thirty-day mortality is modestly lower in patients with NIHSS recorded. Among CMS-linked patients, NIHSS provides excellent discrimination of 30-day mortality among multiple clinically relevant subgroups, and NIHSS as continuous variable results in far better discrimination of 30-day mortality risk than a clinical model not including stroke severity.

Several studies show that NIHSS alone is an extremely robust, if not the strongest, predictor of short-term mortality. A morality-prediction model with NIHSS alone has a c-statistic of 0.82, which is regarded to be genuinely clinically useful.

2b2.2 Analytic Method (*Describe method of validity testing and rationale; if face validity, describe systematic assessment*)**:**
Data from the GWTG-Stroke registry was linked with enrollment files and inpatient claims from the Centers for Medicare & Medicaid Services (CMS) for the period April 1, 2003, through December 31, 2006. Follow-up continued through 2007.

Patients were eligible for inclusion in the GWTG-Stroke registry if they were admitted for acute stroke. Trained hospital personnel ascertained acute ischemic stroke admissions by either prospective clinical identification, retrospective identification using International Classification of Diseases (ICD)-9 discharge codes, or a combination. Patient data abstracted by trained hospital personnel included demographics,

medical history, brain imaging, in-hospital treatment and events, discharge treatment and counseling, mortality, and discharge destination. All patient data were de-identified before submission. All states and regions of the United States were represented and a variety of centers participated, from community hospitals to large tertiary centers. Data on hospital-level characteristics (ie, bed size, academic or nonacademic status, and geographic region) were obtained from the American Hospital Association.

The CMS files (100% Medicare research identifiable files) included data for all fee-for-service Medicare beneficiaries aged =65 years who were hospitalized with a diagnosis of acute stroke (ICD-9- 430.x, 431.x, 433.x, 434.x, and 436.x). We merged patient data in the GWTG registry with Medicare Part A inpatient claims, matching by admission and discharge dates, hospital, date of birth, and sex using methods previously described.From 222278 hospitalizations of patients aged 65 years or older in GWTG-Stroke, we matched 157039 patients (69%) to fee-for-service Medicare claims from 850 hospitals. Patients in Medicare–managed care plans (15% to 25% of the population depending on the region of the country) or other types of insurance are not included in fee-for-service Medicare claims files and therefore cannot be matched. The study population was further confined to patients with acute ischemic stroke (N=101801), first index stroke admission April 2003 or later, and the first index admission (N=94421). Hospitals and patients from centers with <25 ischemic stroke patients entered during the study period were excluded to minimize the likelihood of sampling error. This resulted in a study population of 91134 acute ischemic stroke patients from 625 GWTG-Stroke hospitals.

The NIHSS score was documented in 33770 of these patients (37.1%) from 560 hospitals. For additional stability at the hospital level, we further confined the analysis to hospitals with at least 10 patients with NIHSS documented during the study period. This resulted in a final study population of 33102 patients from 404 participating hospitals. The demographics, clinical characteristics, and geographic distribution of matched and unmatched acute ischemic stroke cases were similar, except fewer patients were enrolled in GWTG-Stroke from the Midwest compared with unmatched ischemic stroke cases. The demographics, clinical characteristics, and geographic distribution of patients with and without NIHSS recorded were similar. The 30-day mortality was modestly lower in the 33102 patients with NIHSS recorded (13.6%) compared with the 57364 patients without NIHSS recorded (14.6%), P<0.0001.

In the harvest data, this variable has a very low missing rate (< 3%.) Thirty-day mortality is modestly lower in patients with NIHSS recorded. Among CMS-linked patients, NIHSS provides excellent discrimination of 30-day mortality among multiple clinically relevant subgroups, and NIHSS as continuous variable results in far better discrimination of 30-day mortality risk than a clinical model not including stroke severity. Several studies show that NIHSS alone is an extremely robust, if not the strongest, predictor of short-term mortality. A morality-prediction model with NIHSS alone has a c-statistic of 0.82, which is regarded to be genuinely clinically useful.[Excerpted from Fonarow GC, Saver JL, Smith EE, Broderick JP, Kleindorfer DO, Sacco RL, et al. Relationship of National Institutes of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries with Acute Ischemic Stroke. J Am Heart Assoc 2012, 1:42-50. Available at http://jaha.ahajournals.org/content/1/1/42]

While we have cited one recent study on NIHSS as predictor of mortality has been confined to patients aged 65 years and older[1], all other studies showing NIHSS is associated with functional status and mortality outcomes were inclusive of ischemic stroke age 18 and older. Below we provide a synopsis of other studies that show the link between functional status and mortality outcomes for patients 18 and older.

One GWTG-Stroke paper shows that NIHSS can be collected in acute ischemic stroke patients of all ages including those <50 years of age and there is not a substantial differences in the % of patients with NIHSS documented by age group (see Table 1, page 882 for more detail) [2]. This study shows that the relationships between age and clinical characteristics, performance measures, and in-hospital outcomes were analyzed in 502 036 ischemic stroke admissions from 1256 hospitals in the Get With the Guidelines–

Stroke program from 2003 to 2009. Data were analyzed by age groups (50, 50 to 59, 60 to 69, 70 to 79, 80 to 89, and 90 years) and with age as a continuous variable. Seven predefined performance measures and 2 summary measures were analyzed. Although modest age-related differences in each individual performance measure were identified, there were substantial temporal improvements in performance measures from 2003 to 2009 in each age group, and many age-related treatment gaps were narrowed or eliminated over time. Older patients were less likely to be discharged home (adjusted odds ratio, 0.69; 95% confidence interval, 0.68 to 0.69) and more likely to die in hospital (adjusted odds ratio, 1.27; 95% confidence interval, 1.25 to 1.29) for each 10-year age increase. In the sensitivity analyses restricted to the cohort of patients in whom NIHSS was recorded, even after adjustment for NIHSS and other covariates, there were similar age-related differences in discharge home and in-hospital mortality. The adjusted OR for in-hospital mortality was 1.15 (95% CI, 1.13 to 1.18; P0.001) per 10-year age increase.

In "A Risk Score for In-Hospital Ischemic Stroke Mortality Derived and Validated within the Get With The Guidelines Stroke Program"[3] researched found that the NIHSS being strongly associated with in-hospital mortality was among patients aged 18 and older. 25% of patients in this GWTG-Stroke study were age 62 or younger. The C statistic for in-hospital mortality in the acute ischemic stroke population age 18 and above is 0.82. Further this risk model showed there was worse prognosis by higher NIHSS score whether the patients was age group <60, 60-70, 70-80, or 80 and higher (for more detail see Figure 3, page 1501).

In addition there is evidence external to GWTG- Stroke data to support NIHSS being recorded in patients age 18 and older.

In one study, a total of 27,671 patients aged 18–80 years treated with IV alteplase within 4.5 hours of symptom onset were enrolled in SITS-ISTR between 2002 and 2010 [3]. The main outcome measures were SICH, death, and favorable functional outcome within 3 months. SICH was considered as local or remote parenchymal hemorrhage type 2,13 on the 22–36 hours posttreatment imaging scan, combined with a neurologic deterioration =4 points on the NIHSS score from baseline, or from the lowest NIHSS value between baseline and 24 hours, or leading to death. Further outcome measures were the proportion of patients with SICH defined, according to NINDS criteria, as any hemorrhage plus a neurologic deterioration (NIHSS score =1) or leading to death within 7 days; the proportion of patients with SICH defined, according to ECASS criteria, as any hemorrhage plus a neurologic deterioration =4 points on the NIHSS from baseline, or from the lowest NIHSS value after baseline to 7 days or leading to death; and no or minimal disability, defined by mRS 0–1 at 3 months. According to NIHSS ranges 0–7, 8–14, and >14, the difference between younger and older patients was highly significant for mortality and functional independence across all NIHSS subgroups (p < 0.0001), while for SICH per SITS-MOST the difference increases over an NIHSS value of 8 (Table 3). Baseline NIHSS, baseline glucose, and signs of infarction in baseline imaging scan were associated with higher mortality and poorer functional outcome. Treatment with IV alteplase is safe in young ischemic stroke patients and they benefit more compared to older patients. We found several factors associated with SICH, mortality, and functional outcome. These can be used to help in the selection of young ischemic stroke patients for thrombolysis.

In another study researchers sough to compare the baseline National Institutes of Health Stroke Scale (NIHSS) score and the Trial of Org 10172 in Acute Stroke Treatment (TOAST) stroke subtype as predictors of outcomes at 7 days and 3 months after ischemic stroke [5]. Using data collected from 1,281 patients enrolled in a clinical trial, subtype of stroke was categorized using the TOAST classification, and neurologic impairment at baseline was quantified using the NIHSS. Outcomes were assessed at 7 days and 3 months using the Barthel Index (BI) and the Glasgow Outcome Scale (GOS). An outcome was rated as excellent if the GOS score was 1 and the BI was 19 or 20 (scale of 0 to 20). Analyses were adjusted for age, sex, race, and history of previous stroke. Researchers founds that the baseline NIHSS score strongly predicted outcome, with one additional point on the NIHSS decreasing the likelihood of excellent outcomes at 7 days by 24% and at 3 months by 17%. At 3 months, excellent outcomes were noted in 46% of patients with NIHSS scores of 7 to 10 and in 23% of patients with scores of 11 to 15. After multivariate adjustment,

lacunar stroke had an odds ratio of 3.1 (95% CI, 1.5 to 6.4) for an excellent outcome at 3 months.

Citations:

[1] Fonarow GC, Saver JL, Smith EE, Broderick JP, Kleindorfer DO, Sacco RL, et al. Relationship of National Institutes of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries with Acute Ischemic Stroke. J Am Heart Assoc 2012, 1:42-50. Available at http://jaha.ahajournals.org/content/1/1/42]) .

[2] Citation: Fonarow GC, Reeves MJ, Zhao X, Olson DM, Smith EE, Saver JL, Schwamm LH; Get With the Guidelines-Stroke Steering Committee and Investigators. Age-related differences in characteristics, performance measures, treatment trends, and outcomes in patients with ischemic stroke. Circulation. 2010 Feb 23;121(7):879-91. Epub 2010 Feb 8.

[3] Citation: Smith EE, Nandavar S, Dai D, Olson D, Reeves M, Saver J, Hernandez A, Peterson E, Fonarow G, Schwamm L. A Risk Score for In-Hospital Ischemic Stroke Mortality Derived and Validated within the Get With The Guidelines Stroke Program. Circulation.2010; 122: 1496-1504 Available at: circ.ahajournals.org/content/122/15/1496.full

[4] Toni D, Ahmed N, Anzini A, Lorenzano S, Brozman M, Kaste M, Mikulik R, Putaala J, Wahlgren N; SITS investigators. Intravenous thrombolysis in young stroke patients: results from the SITS-ISTR. Neurology. 2012 Mar 20;78(12):880-7. Epub 2012 Mar 7. Available at http://www.neurology.org/content/78/12/880.abstract]

[5] Citation: Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, Woolson RF, Hansen MD. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999 Jul 13;53(1):126-31. Available at: http://www.neurology.org/content/53/1/126.full

2b2.3 Testing Results (*Statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment*):
The measure has been empirically tested and applied for over 9 years and 1600 plus US hospitals in GWTG Stroke and directly demonstrated to be useful in improving patient outcomes, highly interpretable and actionable. The measure design has been demonstrated to have a denominator precisely defined, numerator precisely defined. The measure has been shown to have face, content, and construct validity. Further peer reviewed published data have established reliability, measure implementation feasibility has been demonstrated on a national scale at reasonable effort, reasonable cost, reasonable time period for collection, and have been assessed by explicit, predefined criteria for inclusion in a stroke performance measurement set.

A review of the relevant evidence and guidelines and expert panel consensus process resulted in the conclusion NIHSS is important when assessing initial stroke severity of a patient, as this has been shown to predict mortality in acute ischemic stroke.

For NIHSS measure, 80% of the expert panel strongly agreed with the value of this measure with 20% agree. This group is comprised of experts in cerebrovascular and cardiovascular disease who hold prominent roles in societies. Members are nominated by peers across the United States to participate, and member selection includes approval of leadership of the American Heart Association based on the nominee's clinical expertise. Several indviduals who serve on the Get With The Guidelines-Stroke Clinical Workgroup and Executive Committee provide leadership in outcomes research. A list of those volunteers inovolved is included in the "additional" section of this form.

POTENTIAL THREATS TO VALIDITY. (*All potential threats to validity were appropriately tested with*

adequate results.)

2b3. Measure Exclusions. (*Exclusions were supported by the clinical evidence in 1c or appropriately tested with results demonstrating the need to specify them.*)

2b3.1 Data/Sample for analysis of exclusions (*Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*)**:**
Data for exclusion is compared based on results from the derivation cohort from 2011 with data from 2010.

Exclusion criteria for this measure at the level of hospital stays, patients, and facilities is presented in the accompanying output. The starting sample was the total number of records submitted to Get With The Guidelines-Stroke during the associated year.

2b3.2 Analytic Method (*Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference*)**:**
Frequency of exclusion coding: Beginning with the overall sample, the total number and associated of hospital stays, patients, or facilities impacted by each sequential exclusion was calculated until the final sample was reached. Each subsequent exclusion was calculated from the sample remaining.

2b3.3 Results (*Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses*):
Rates of exclusions are essentially the same across both the derivation (2011 – beginning hospital stays n= 353,731; ending hospital stays n= 258,157 (representing 73% of the original sample)) and the validation cohorts (2010 – beginning hospital stays n= 364,275; ending hospital stays n= 262,109 (representing 72% of the original sample)).

2b4. Risk Adjustment Strategy. (*For outcome measures, adjustment for differences in case mix (severity) across measured entities was appropriately tested with adequate results.*)

2b4.1 Data/Sample (*Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*)**:**
Not applicable, no risk adjustment performed.

2b4.2 Analytic Method (*Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables*)**:**
Not applicable, no risk adjustment performed.

2b4.3 Testing Results (*Statistical risk model: Provide quantitative assessment of relative contribution of model risk factors; risk model performance metrics including cross-validation discrimination and calibration statistics, calibration curve and risk decile plot, and assessment of adequacy in the context of norms for risk models. Risk stratification: Provide quantitative assessment of relationship of risk factors to the outcome and differences in outcomes among the strata*)**:**
Not applicable, no risk adjustment performed.

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: Not applicable.

2b5. Identification of Meaningful Differences in Performance. (*The performance measure scores were appropriately analyzed and discriminated meaningful differences in quality.*)

2b5.1 Data/Sample (*Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*)**:**
Reliability was established by validating the derivation cohort from 2011 with data from 2010. After exclusions, a total of 258,157 hospital stays (representing 221,908 patients) were submitted by 1,535 facilities were eligible for this measure in the derivation cohort.

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

2b5.2 Analytic Method *(Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance)***:**
Reliability was established by comparing performance results from the derivation cohort from 2011 with data from 2010.

2b5.3 Results *(Provide measure performance results/scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:
Compared to the validation cohort, the percent of patients for which an NIHSS was performed was consistently higher. At the hospital level, the mean of patients in 2011 having had an NIHSS administered was approximately 58%, compared to the 2010 rate of 53%. Similarly, the median was 68% to 59%, respectively. For the derivation and the validation years, the lowest decile was 3% and 0%, respectively. The highest decile was 95% for the 2011 derivation cohort and 94% for the 2010 validation cohort. These increases are what one would normally expect given hospitals local efforts to improve these rates over time and the increasing importance ascribed to this measure of stroke severity.

2b6. Comparability of Multiple Data Sources/Methods. *(If specified for more than one data source, the various approaches result in comparable scores.)*

2b6.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)***:**
This section is not applicable. Only the Get With The Guidelines-Stroke database was used as the source of testing for this measure.

2b6.2 Analytic Method *(Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure)***:**
This section is not applicable. Only the Get With The Guidelines-Stroke database was used as the source of testing for this measure.

2b6.3 Testing Results *(Provide statistical results, e.g., correlation statistics, comparison of rankings; assessment of adequacy in the context of norms for the test conducted)***:**
This section is not applicable. Only the Get With The Guidelines-Stroke database was used as the source of testing for this measure.

2c. Disparities in Care: H⊙ M⊙ L⊙ I ⊙ NA⊙ *(If applicable, the measure specifications allow identification of disparities.)*

2c.1 If measure is stratified for disparities, provide stratified results *(Scores by stratified categories/cohorts)***:** Across all races, Hispanics were consistently more likely to be administered the NIHSS. Similar to trends in performance noted above, these disparities were more striking in the derivation cohort (Hispanics: mean 64%; median 77%). In looking at quartiles of each race by across hospital patient submissions, in both the derivation and the validation cohorts, races were consistently more likely to be administered the NIHSS in the 2nd and 3rd quartiles with drop-off in both the fewest and highest quartiles of patient populations represented. Once again, however, the exception is the Hispanic subgroup of the derivation cohort that demonstrated a more linear increase in NIHSS administration across the 4 four quartiles with essentially equivalent rates for the 3rd and 4th quartiles.

2c.2 If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:
This measure is a process of care measure. It has been well documented in the literature that processes of care differ amongst racial groups.

2.1-2.3 Supplemental Testing Methodology Information:
Attachment
NIHSS_Recorded_2010_2011.pdf

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

Steering Committee: Overall, was the criterion, *Scientific Acceptability of Measure Properties*, met?
(*Reliability and Validity must be rated moderate or high*) Yes⦿ No⦿
Provide rationale based on specific subcriteria:

If the Committee votes No, STOP

3. USABILITY

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

C.1 Intended Actual/Planned Use (*Check all the planned uses for which the measure is intended*)**:**
Professional Certification or Recognition Program, Public Reporting, Quality Improvement (Internal to the specific organization), Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

3.1 Current Use (*Check all that apply; for any that are checked, provide the specific program information in the following questions*): Public Reporting, Professional Certification or Recognition Program, Quality Improvement with Benchmarking (external benchmarking to multiple organizations), Quality Improvement (Internal to the specific organization)

3a. Usefulness for Public Reporting: H⦿ M⦿ L⦿ I ⦿
(*The measure is meaningful, understandable and useful for public reporting.*)

3a.1. Use in Public Reporting - disclosure of performance results to the public at large (*If used in a public reporting program, provide name of program(s), locations, Web page URL(s)*). If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: [***For Maintenance*** *– If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.*]
This measure is not yet used in any public reporting initiative. The AHA/ASA will explore how to get this measure adopted in government programs and would thus provide information about clinician participation to the public. The goal of all performance measures is to link processes of care to meaningful outcomes. As it is an evolving process, we are evaluating public reporting options.

3a.2.**Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting**. If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: The NIHSS is a good indicator of the risk for 30 day mortality. By increasing the number of hospitals that routinely perform the NIHSS scale can help in the discrimination of mortality for patients. However to do so, will require a greater reporting rate of the NIHSS.

The reporting of NIHSS can help hospitals and providers to have a better understanding of the importance of this functional status assessment in patients. Furthermore, inclusion of this measure in a public reporting program has the potential to increase awareness of the importance of this measure in possible reducing 30 day mortality in acute ischemic stroke patients.

3.2 Use for other Accountability Functions (payment, certification, accreditation). If used in a public accountability program, provide name of program(s), locations, Web page URL(s): This measure is also used in the Get With The Guidelines- Stroke Plus Awards. To read more about our plus awards recognition program go to

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

http://www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelinesHFStroke/GetWithTheGuidelinesStrokeHomePage/Get-With-The-Guidelines-Stroke-Recognition-Criteria_UCM_310337_Article.jsp. Click on Gold Plus Quality Awards or Silver Plus Quality Awards.

3b. Usefulness for Quality Improvement: H⚫ M⚪ L⚪ I ⚪
(*The measure is meaningful, understandable and useful for quality improvement.*)

3b.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):
[For **Maintenance** – *If not used for QI, indicate the reasons and describe progress toward using performance results for improvement***]**.
Get With the Guidelines-Stroke is the American Heart Association's collaborative performance improvement program, demonstrated to improve adherence to evidence-based care of patients hospitalized with stroke. This program includes the measure we have submitted and is used in over 1,600 hospitals.

To find out general information on Get With The Gudelines-Stroke, please go to:
http://www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelinesHFStroke/Get-With-The-Guidelines-Stroke-Home-Page_UCM_306098_SubHomePage.jsp .

Hospitals teams that participate actively and consistently in Get With The Guidelines-Stroke get rewarded with public recognition by providing patients and stakeholders with tangible evidence of their commitment to improving quality care. Silver, Gold, Silver Plus and Gold Plus award-winning Get With The Guidelines-Stroke hospitals are honored at national recognition events during the International Stroke Conference and listed by name in advertisements that appear annually in the journal Stroke and in the "Best Hospitals" issue of U.S. News & World Report. To read more about our plus awards recognition program go to http://www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelinesHFStroke/GetWithTheGuidelinesStrokeHomePage/Get-With-The-Guidelines-Stroke-Recognition-Criteria_UCM_310337_Article.jsp. Click on Gold Plus Quality Awards or Silver Plus Quality Awards.

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated *(e.g., QI initiative),* describe the data, method and results:
The measure is use by over 1,600 US hospitals participating in Get With The Guidelines- Stroke which collects this measure. The NIHSS measure is included in the GWTG Stroke plus awards.

To read more about our plus awards recognition program go to
http://www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelinesHFStroke/GetWithTheGuidelinesStrokeHomePage/Get-With-The-Guidelines-Stroke-Recognition-Criteria_UCM_310337_Article.jsp. Click on Gold Plus Quality Awards or Silver Plus Quality Awards.

Overall, to what extent was the criterion, *Usability*, met? H⚫ M⚪ L⚪ I ⚪
Provide rationale based on specific subcriteria:

4. FEASIBILITY

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (**evaluation criteria**)

4a. Data Generated as a Byproduct of Care Processes: H⚫ M⚪ L⚪ I ⚪

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*
Data used in the measure are**:**
Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

4b. Electronic Sources: H⦿ M⦿ L⦿ I ⦿

4b.1 Are the data elements needed for the measure as specified available electronically *(Elements that are needed to compute measure scores are in defined, computer-readable fields)*: Some data elements are in electronic sources

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources: There are clinical exclusion criteria that may not be part of the standard electronic data sets included in the CCR/CCD. The AHA/ASA hopes to work to attempt to get code sets developed to address exclusions, and will work to identify appropriate codes that would be needed and work with the appropriate organization(s) to create appropriate additional necessary codes for exclusions.

4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H⦿ M⦿ L⦿ I ⦿

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:

There are current risk-standardization and risk-adjustment models that fail to capture the complexity of stroke severity on admission and may have the unintended consequence of encouraging clinicians to avoid treatment of the most high-risk patients who are most likely to receive treatment.

The NIH Stroke Scale is a well-validated tool to assess initial stroke severity. (1,2)   A recent paper in Stroke highlights the interest in reporting outcomes for Medicare patients hospitalized with acute ischemic stroke.(3) An evidence base exists for reported acute ischemic stroke severity on admission being linked to outcomes.(4-6) There are no reported unintended consequences.

Fonarow and colleagues reported a c-statistic of 0.82 for NIHSS score alone predicting mortality using data from 404 hospitals and 33,102 Medicare beneficiaries. A total of 4,496 deaths were reported in the first 30 days after stroke and the those scoring 22-42 on the 42 point scale represented more than half the deaths (53.3.%). (7)

Citations:

1. Brott T, Adams HP Jr, Olinger CP, Marler JR, Barsan WG, Biller J, Spilker J, Holleran R, Eberle R, Herztberg V, Rorick M, Moonaw CJ, Walker M. Measurement of acute cerebral infarction: a clinical examination scale. Stroke. 1989;20:864–870. Available at: http://stroke.ahajournals.org/content/20/7/864.abstract

2.Goldstein LB, Bertels C, Davis JN. Interrater reliability of the NIH Stroke Scale. Arch Neurol. 1989;46:660–662. Available at: http://archneur.ama-assn.org/cgi/reprint/46/6/660

3. Lichtman JH, Leifheit-Limson EC, Jones SB, Watanabe E, Bernheim SM, Phipps MS, Bhat KR, Savage SV, Goldstein LB. Predictors of hospital readmission after stroke: a systematic review. Stroke. 2010;41:2525–2533. Available at: http://stroke.ahajournals.org/content/41/11/2525.full

4. Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, Woolson RF, Hansen MD. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: a report of the Trial of ORG 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999;53:126–131. Available at: http://www.neurology.org/content/53/1/126.full

5. Dhamoon MS, Moon YP, Paik MC, Boden-Albala B, Rundek T, Sacco RL, Elkind MS. Long-term functional recovery after first ischemic stroke: the Northern Manhattan Study. Stroke. 2009;40:2805–2811.

Available at: http://stroke.ahajournals.org/content/early/2009/06/25/STROKEAHA.109.549576.full.pdf+html

6. Johnston KC, Connors AF Jr, Wagner DP, Knaus WA, Wang X, Haley EC Jr. A Predictive Risk Model for Outcomes of Ischemic Stroke. Stroke. 2000;31:448–455. Available at: http://stroke.ahajournals.org/content/31/2/448.full.pdf

7. Relationship of National Institutes of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries With Acute Ischemic Stroke. Fonarow GC,Saver JL, Smith EE, Broderick JP, Kleindorfer DO, Sacco RL, Pan W, Olson DM, Hernandez AF, Peterson ED, Schwamm LH. J Am Heart Assoc. 2012;1:42-50. Available at: http://jaha.ahajournals.org/content/1/1/42.full

---

4d. Data Collection Strategy/Implementation: H⬤ M⬤ L⬤ I ⬤

A.2 Please check if either of the following apply (*regarding proprietary measures*)**:**
4d.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues (*e.g., fees for use of proprietary measures*)**:**

Get With The Guidelines-Stroke has been tracking the percent of patients with a documented NIHSS since 2003. Hospitals have used this metric for internal quality improvement, state reporting (in some states, this is a metric for stroke center designation), and to qualify for GWTG PLUS Awards. Since 2003, both the measure logic and data elements used in the measure have undergone minor refinements as a result of data analysis and hospital feedback. The measure originally included ischemic stroke, hemorrhagic stroke, stroke not otherwise specified, but hemorrhagic stroke patients were removed to reflect the appropriate use of alternative scales in these patients. We have added the exclusion of patients whose symptoms had resolved upon presentation, as an NIHSS would not be performed in most of these cases as these patients have no deficits.

This measure requires the collection of only two additional data elements beyond what is collected for STK measures: Initial NIHSS Score and whether Stroke Symptoms Resolved at Time of Presentation. These data elements should require minimal additional work on the part of the hospital as the performance of an NIHSS or comprehensive neurological exam are part of a standard stroke protocol for most facilities. Furthermore, stroke severity (too mild or too severe) is used as a consideration for tPA eligibility. We would expect that stroke severity would be documented in the chart using the NIHSS score or comprehensive neurological exam findings.

Overall, to what extent was the criterion, *Feasibility*, met? H⬤ M⬤ L⬤ I ⬤
Provide rationale based on specific subcriteria:

OVERALL SUITABILITY FOR ENDORSEMENT

Does the measure meet all the NQF criteria for endorsement? Yes⬤ No⬤
Rationale:

If the Committee votes No, STOP.
If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

---

5. COMPARISON TO RELATED AND COMPETING MEASURES

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

---

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

best measure before a final recommendation is made.

5.1 If there are related measures *(either same measure focus or target population)* **or competing measures** *(both the same measure focus and same target population)*, **list the NQF # and title of all related and/or competing measures:**

5a. Harmonization

5a.1 If this measure has EITHER the same measure focus OR the same target population as NQF-endorsed measure(s): Are the measure specifications completely harmonized?

5a.2 If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden:

5b. Competing Measure(s)

5b.1 If this measure has both the same measure focus and the same target population as NQF-endorsed measure(s):
Describe why this measure is superior to competing measures *(e.g., a more valid or efficient way to measure quality)*; **OR provide a rationale for the additive value of endorsing an additional measure.** *(Provide analyses when possible)*:

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner): American Heart Association/American Stroke Association, 7272 Greenville Avenue, Dallas, Texas, 75231

Co.2 Point of Contact: Penelope, Solis, JD, penelope.solis@heart.org, 202-423-3124-

Co.3 Measure Developer if different from Measure Steward: American Heart Association/American Stroke Association, 7272 Greenville Avenue, Dallas, Texas, 75231

Co.4 Point of Contact: Penelope, Solis, JD, penelope.solis@heart.org, 202-423-3124-

Co.5 Submitter: Penelope, Solis, JD, penelope.solis@heart.org, 202-423-3124-, American Heart Association/American Stroke Association

Co.6 Additional organizations that sponsored/participated in measure development: Not Applicable.

Co.7 Public Contact: Penelope, Solis, JD, penelope.solis@heart.org, 202-423-3124-, American Heart Association/American Stroke Association

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
The below volunteers are part of the Get With The Guidelines-Stroke measures workgroup and are responsible for developing and maintaining measures included in the Get With The Guidelines-Stroke module.

*Eric E. Smith, MD, MPH, FRCPC

Chair
Assistant Neurologist
Massachusetts General Hospital

*Lee H. Schwamm, MD, FAHA
Vice Chairman of the Neurology Dept
Massachusetts General Hospital

*Gregg C. Fonarow, M.D., FACC
Professor of Medicine
Director, Ahmanson-UCLA Cardiomyopathy Center
Co-Director, UCLA Preventative Cardiology Program

Jeff Saver, MD, FAHA, FAAN
Professor of Neurology
Geffen School of Medicine at UCLA

Mathew Reeves, PhD, DVM
Associate Professor
Department of Epidemiology
Michigan State University

David Tong MD FAHA
Medical Director, CPMC Comprehensive Stroke Care Center
Director, CPMC Center for Stroke Research (CCSR)

Scott Kasner, MD, MSCE, FAHA
Professor of Neurology
Director, Comprehensive Stroke Center
University of Pennsylvania
Medical Center

Measures were reviewed by the GWTG- Exec Committee which includes the below volunteers as well as those above denoted with an asterisk:

Paul Heidenreich, MD,MS
Associate Professor of Medicine
Stanford University
VA Palo Alto Medical Center

Robert Berg, MD, FAAP, FAHA
Professor & Division Chief
Critical Care Medicine
Children's Hospital Philadelphia

Eric D. Peterson, MD, MPH, FAHA, FACC
Professor of Medicine
Vice Chair for Quality
Duke University Medical Center
Associate Director and Director of CV Research
Duke Clinical Research Institute

Adrian Hernandez, MD
Duke University Medical Cntr

Deepak L. Bhatt, MD, MPH, FAHA, FACC, FSCAI
Chief of Cardiology, VA Boston Healthcare System
Director, Integrated Interventional Cardiovascular Program, Brigham and Women's Hospital & VA Boston Healthcare System
Associate Professor of Medicine, Harvard Medical School

Ad.2 If adapted, provide title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward: Not Applicable.

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.3 Year the measure was first released: 2003
Ad.4 Month and Year of most recent revision: 04, 2012
Ad.5 What is your frequency for review/update of this measure? Annual
Ad.6 When is the next scheduled review/update for this measure? 04, 2013

Ad.7 Copyright statement: © 2012 American Heart Association/American Stroke Association. All Rights Reserved.

Ad.8 Disclaimers:

Ad.9 Additional Information/Comments: To read additional guidelines which support NIHSS being performed, please read the VA/DOD Stroke Rehabilitation Guidelines. Accessible at: http://www.healthquality.va.gov/stroke/stroke_full_221.pdf

Date of Submission (*MM/DD/YY*): 05/03/2012