Journal of the American Heart Association

BRIEF COMMUNICATION

Reliability of Past Medical History in a Single Hospital Participating in Get With The Guidelines-Stroke Registry

Christopher G. Favilla, MD; Alice F. Ford, MD, PhD; Ossama Khazaal, MD; Daniel Cristancho, MD; Emily Grodinsky, MD; Judy Dawod, MD; Scott E. Kasner, MD

BACKGROUND: The GWTG (Get With The Guidelines)-Stroke registry supports clinical research and quality improvement projects that often rely on past medical history elements, the reliability of which remains largely unknown. Here, we evaluated the reliability of specific past medical history elements in a local GWTG–Stroke data set, with particular attention to calculating the CHA_2DS_2-VASc score.

METHODS AND RESULTS: A single-center cohort was identified by querying the Hospital of the University of Pennsylvania's GWTG IQVIA Registry Platform for patients admitted with acute ischemic stroke between January 2017 and December 2020, with a previously known history of atrial fibrillation. Demographics and previously known medical history elements were retrieved from the registry to calculate the CHA_2DS_2-VASc score. Five neurologists abstracted the same medical history elements from the health records. The κ statistics quantified the reliability of medical history elements and CHA_2DS_2-VASc score. Four hundred fifty-three patients with acute ischemic stroke and previously known atrial fibrillation were included in the cohort. In comparison with manual reabstraction, registry-based medical history elements were only moderately reliable: congestive heart failure (κ=0.53), hypertension (κ=0.42), diabetes (κ=0.80), prior stroke (κ=0.45), and vascular disease (κ=0.48). However, leveraging these variables to calculate the CHA_2DS_2-VASc score was more reliable (κ=0.73).

CONCLUSIONS: Previously known medical history elements in the GWTG-Stroke registry were only modestly reliable in this single-center study, suggesting caution should be exercised when relying on any individual history elements in registry-based research. Combining these variables to calculate the CHA_2DS_2-VASc score was somewhat more reliable. Multicenter data are needed before assuming generalizability.

Key Words: atrial fibrillation ■ Get With The Guidelines ■ registry ■ stroke

The GWTG (Get With The Guidelines)-Stroke registry is an American Heart Association (AHA)-supported quality improvement program aimed at promoting consistent adherence to evidence-based treatment guidelines.[1–3] The GWTG-Stroke registry also supports a wide range of clinical research and quality improvement projects at the local and national level. The registry relies on local data abstraction and completion of a case record form, the reliability of which has largely been established by comparing registry data with manual chart review.[4] However, past medical history warrants unique consideration, because unlike demographics, time points, or medication administration, the key elements of the medical history rely on a provider's ability to obtain and document a thorough history. Moreover, given the multidisciplinary nature of stroke care, there are often multiple providers writing notes on each patient, which could also lead to inconsistency. Reliability of these data is critical, because registry-based analyses frequently target a specific

Correspondence to: Christopher G. Favilla, MD, Department of Neurology, University of Pennsylvania, 3400 Spruce Street, 3 Gates, Philadelphia, PA 19146. Email: christopher.favilla@pennmedicine.upenn.edu

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.122.025308

For Sources of Funding and Disclosures, see page 5.

JAHA is available at: www.ahajournals.org/journal/jaha

population based on the presence of a specific vascular risk factor.[5–8]

Past medical history fields can also be leveraged to calculate risk-stratification scores, such as the CHA_2DS_2-VASc score, which estimates the annualized risk of stroke in patients with atrial fibrillation (AF).[9,10] Although the CHA_2DS_2-VASc score is not directly collected in GWTG-Stroke, registry data can be leveraged to calculate the CHA_2DS_2-VASc score in future studies, including those exploring why some patients with AF are not on anticoagulation at the time of their stroke.[11] Here, we evaluated the reliability of specific past medical history elements in the local GWTG-Stroke data set, with particular attention to calculating the CHA_2DS_2-VASc score.

METHODS

This was a single-center, retrospective cohort study approved by the University of Pennsylvania Institutional Review Board. A waiver of informed consent was granted. The data that support the findings are available from the corresponding author upon reasonable request.

GWTG-Stroke Database and Generation of CHA_2DS_2-VASc Score

The authors queried the Hospital of the University of Pennsylvania's local GWTG-Stroke IQVIA Registry Platform to identify all patients admitted between January 2017 and December 2020, with a discharge diagnosis of acute ischemic stroke and AF that was known before the stroke. Patients with newly detected AF during the stroke hospitalization were excluded. Admission date, patient demographics, previously known medical history, and initial National Institutes of Health Stroke Scale score were retrieved from the IQVIA Registry Platform. Vascular disease, as defined by the CHA_2DS_2-VASc score, is not explicitly captured by the GWTG-Stroke case record form, so it was scored based on the presence of at least 1 of the following diagnoses before the index stroke: coronary artery disease/prior myocardial infarction, carotid stenosis, or peripheral vascular disease. This work represents the authors' independent analysis of local data gathered using the AHA GWTG IQVIA Registry Platform but is not an analysis of the national GWTG data set and does not represent findings from the AHA GWTG National Program. Registry data at the University of Pennsylvania are populated by a team of trained data abstractors, led by a nurse with >15 years of data abstraction experience. The lead abstractor was responsible for the majority of charts but was supplemented by 2 neuroscience nurses and 2 stroke research coordinators. All GWTG abstractors completed standard training. The lead abstractor routinely reviewed charts that others abstracted to ensure interrater reliability was appropriately maintained. Standard instructions for GWTG abstraction indicate that medical history elements should be abstracted from a range of documentation during the index admission. The Hospital of the University of Pennsylvania GWTG database has been maintained for more >15 years.

Chart Abstraction

Manual chart review was performed for each patient identified by the above query. For this reabstraction process, review was limited to documentation during the stroke hospitalization to mirror the GWTG data abstraction. Abstraction instructions are available in Data S1 and closely replicated the standardized instructions used by the trained GTWG abstractors. Components of the previously known medical history were abstracted, focusing on components of the CHA_2DS_2-VASc score: congestive heart failure, hypertension, diabetes, prior stroke/transient ischemic attack/thromboembolism, and vascular disease. As per the CHA_2DS_2-VASc score, vascular disease was defined as a previous myocardial infarction, peripheral arterial disease, or aortic plaque.[10] Previously known elements of the medical history were defined as conditions known to exist before the admission for the acute stroke. Conditions that are newly diagnosed during the stroke hospitalization are not considered part of the previously known medical history but are separately categorized as newly diagnosed. Separately, coding newly diagnosed variables were used to calculate the discharge CHA_2DS_2-VASc score. Reabstraction was divided evenly among 5 neurologists.

Statistical Analysis

Interrater reliability of the registry-calculated and reabstracted CHA_2DS_2-VASc score was assessed by quadratic-weighted κ statistic. A Bland-Altman plot depicted the agreement between the 2 CHA_2DS_2-VASc scoring methods. A stratified analysis was performed to compare the 5 physicians performing manual reabstraction. Dichotomized CHA_2DS_2-VASc score (<2 versus ≥ 2) and individual past medical history elements were compared by Cohen κ. A McNemar test assessed if a medical history element was more likely to be reported by one particular scoring method. All tests were evaluated at a significance level of $P<0.05$. Statistical analyses were performed in Stata 15.1 (StataCorp, College Station, TX).

In considering sample size, it was estimated that 80% to 90% would have a CHA_2DS_2-VASc score of ≥ 2. A 10% difference between the 2 scoring techniques was conservatively estimated to ensure a sufficiently large sample size was selected. Setting α to 0.05, a

sample of 361 would be adequate to detect a κ of 0.70. At least 100 patients met eligibility criteria annually, so a 4-year sample was selected.

RESULTS

Four hundred fifty-seven patients were identified as having a discharge diagnosis of acute ischemic stroke and a previously known history of AF. Four patients were excluded after manual review; 1 had an incorrect diagnosis of stroke, and 3 had inaccessible medical records. The final cohort consisted of 453 patients. Mean age was 75 (±12) years. The cohort was 49% women. Mean admission National Institutes of Health Stroke Scale score was 13 (±9). Fifty-eight percent were White, 28% were Black, 3% Asian, and 11% undetermined. There were no missing data for age, sex, and previously known medical history. Three percent were missing National Institutes of Health Stroke Scale scores. Table 1 summarizes the medical history and total CHA_2DS_2-VASc score based on GWTG registry abstraction and manual reabstraction.

The Figure shows the concordance of registry-calculated and reabstracted CHA_2DS_2-VASc scores. Quadratic-weighted κ was 0.73. The κ was similar across the 5 reviewers (Figure S1). The scatterplot (Figure – Panel A) reveals a linear relationship between the 2 scoring techniques (r=0.73). Notably, at the low end of the spectrum, the registry-calculated technique may overestimate the score compared with the reabstracted score, whereas at the high end of the spectrum, the registry-calculated technique may underestimate the score. A Bland-Altman analysis (Figure – Panel B) shows that the overall mean difference between the 2 scoring techniques was 0.20 (nearly 0), suggesting high concordance overall, but the 95% CIs for agreement are broad (95% CI, –2.8 to 2.4), indicating substantial disagreement for individual patients. When dichotomizing the pre-admission total CHA_2DS_2-VASc score as <2 versus ≥2, there was 95% agreement between the scoring methods, and the κ was 0.69.

Despite reliability of the overall CHA_2DS_2-VASc score, individual past medical history elements were less reliably reported (Table 1). A significant minority of cases had a discrepancy with respect to an individual essential element. Previously known diagnoses of congestive heart failure, hypertension, and diabetes were underreported in the GWTG-Stroke registry (Table 2). Prior stroke and vascular disease were common sources of disagreement (24% and 23%, respectively). Prior stroke discrepancies were fairly well balanced between the 2 scoring techniques. Vascular disease may have been scored more often by the GWTG scoring technique but this imbalance was nonsignificant (P=0.09).

During manual reabstraction, new diagnoses (during the stroke admission) of vascular risk factors were recorded. New diagnoses were not included in the previously known medical history as reported above. A new diagnosis of congestive heart failure was made in 15% of patients, diabetes in 8%, hypertension in 10%, and vascular disease in 7%. The mean total CHA_2DS_2-VASc score was 4.3 on admission and 5.8 on discharge (35% of the study population had a prior history of stroke before the admission).

DISCUSSION

Clinical registries, such as GWTG-Stroke, provide the infrastructure to support unique clinical research and quality improvement efforts, but data validity is critical. The scale of data collection in GWTG-Stroke presents challenges at each participating hospital. In this single-center study, the reliability of most individual past medical history elements was modest at best. The reliability of the registry-calculated CHA_2DS_2-VASc score was somewhat more substantial. It may be reasonable to leverage these registry data to calculate the CHA_2DS_2-VASc score if applicable to a given quality improvement or clinical research project, but these findings also suggest exercising caution when relying heavily on certain individual components of the previously known medical history. These findings reflect local data gathered using the AHA GWTG IQVIA Registry Platform but not national-level data.

The overall CHA_2DS_2-VASc score may be more reliable than the dichotomized score (κ=0.73 and 0.68, respectively). This could be explained by a larger number of data elements in the total score, but comparing

Table 1. Past Medical History: Comparing the GWTG-Stroke Registry and Manual Abstraction

Medical history element	Based on the GWTG	Based on manual reabstraction	Agreement	κ
Congestive heart failure	23%	32%	81%	0.53
Hypertension	75%	82%	81%	0.42
Diabetes	29%	32%	92%	0.80
Prior stroke	32%	33%	76%	0.45
Vascular disease	33%	29%	78%	0.48

Agreement was assessed by Cohen's κ. GWTG indicates Get With The Guidelines.

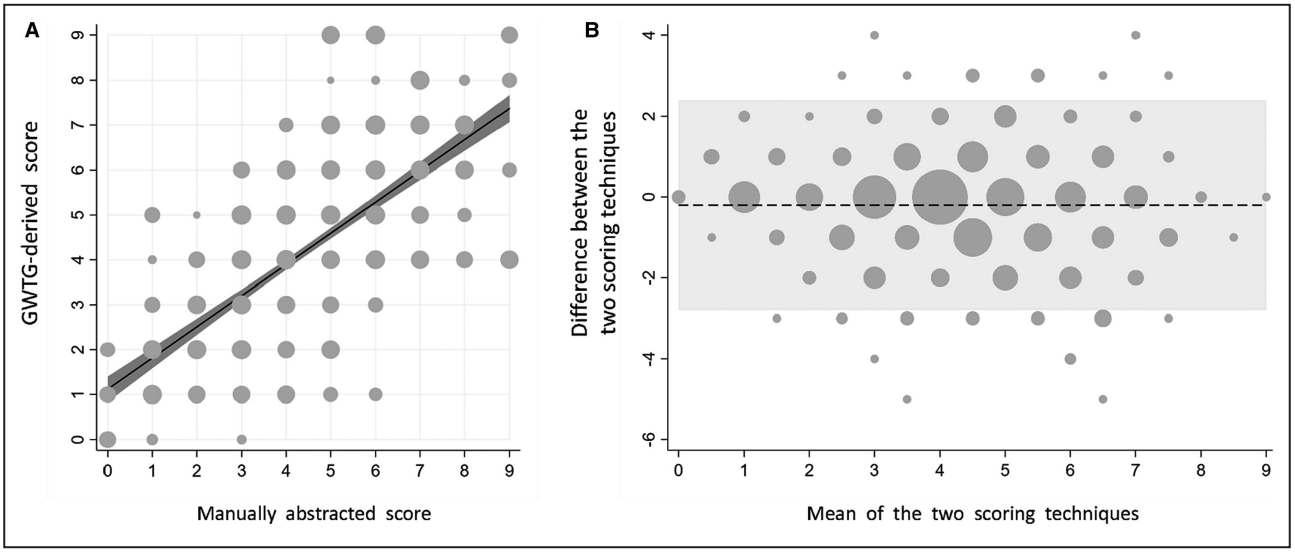


Figure. Reliability of calculating the CHA_2DS_2-VASc from the GWTG (Get With The Guidelines)-Stroke registry.
A, The scatterplot depicts the CHA_2DS_2-VASc score based on manual reabstraction (*x* axis) and calculation by GWTG-Stroke registry data (*y* axis). Quadratic weighted κ=0.73. The solid black line represents the linear fit line (*R*=0.73) and the shaded area represents the 95% CIs. **B**, A Bland-Altman plot shows that the overall mean difference between the 2 techniques was 0.20 (dashed line). The 95% CIs for agreement extend from −2.8 to 2.4 (shaded area). In both panels, the dot size reflects the number of patients at a given point.

κ statistics is limited because the overall score was assessed by quadratic-weighted κ, and the majority of discrepant scores only differed by 1. It is important to note the CHA_2DS_2-VASc score discussed here reflects the score at the time of hospital admission, and therefore does not reflect new diagnoses made during the stroke hospitalization. The CHA_2DS_2-VASc score before the stroke admission may be important when assessing why some patients with AF are not on anticoagulation at the time of their stroke.[11]

In a national-level evaluation of data accuracy, Xian et al reported excellent reliability of GWTG-Stroke registry data but did not report individual history elements.[4] Many registry data fields, such as demographics, time points, and laboratory values have little uncertainty during abstraction. The presence or absence of a medical history element seems straightforward at first pass but in practice relies heavily on a clinician's ability to obtain a thorough history and document accordingly. Moreover, the GWTG-Stroke registry abstractors and the manual chart reviewers in this study had access to the same medical records and were given the same instructions for abstraction, so these results may reveal a discrepancy in how clinical documentation is interpreted. The high rate of discrepant past medical history elements should be considered when future studies rely on these variables. Initial review of the summary statistics suggested prior stroke was relatively well balanced between the 2 scoring techniques. However, this apparent balance was a result of discrepancies being common with both scoring techniques, reflected in a relatively weak κ statistic. Heart failure, hypertension, and diabetes were more likely to be scored by manual reabstraction, which raises the possibility that physicians performing reabstraction were more thorough than GWTG abstractors. Alternatively, physicians performing reabstraction may have been more comfortable interpreting unclear provider documentation. Vascular disease, on the other hand, was less likely to be scored by manual reabstraction, which may reflect an overly sensitive strategy for scoring vascular disease based on GWTG data. Multicenter data are required to determine the generalizability of these finding.

Table 2. Discrepant Past Medical History Elements

Medical history element	Only noted via GWTG	Only noted via reabstraction	*P* value
Congestive heart failure	24 (5%)	62 (14%)	<0.001
Hypertension	28 (6%)	60 (13%)	0.0006
Diabetes	12 (3%)	27 (6%)	0.02
Prior stroke	49 (11%)	61 (13%)	0.25
Vascular disease	60 (13%)	42 (9%)	0.08

P values calculated by McNemar test. GWTG indicates Get With The Guidelines.

These results should be interpreted within the context of several limitations. Most notably, as a single-center study, these findings may not reflect data reliability at other centers. The retrospective nature of the study may be perceived as a limitation, but this more accurately reflects real-world use of the registry data. Furthermore, this study did not explore all medical history elements on

the case record form, but instead focused on elements that contribute to the CHA_2DS_2-VASc score. The authors have no reason to suspect the variables reviewed here are more or less reliable than other medical history variables not reviewed. The manner in which registry data are leveraged to assess vascular disease presents a limitation because the CHA_2DS_2-VASc score defines vascular disease as the presence of prior myocardial infarction, peripheral arterial disease, carotid stenosis, or complex aortic plaque. Here, vascular disease was more likely to be scored in the GWTG-based score, but the reliability of vascular disease was similar to that of more straightforward variables such as hypertension, heart failure, and prior stroke. Assessing reliability or validity is also limited by the gold standard to which the comparison is made. Here, the manual chart review might be perceived as the gold standard, but the registry is similarly populated by manual chart review. Both methods of assessment may suffer from similar limitations, so several reviewers were used in this study to minimize the possibility of idiosyncratic tendencies with record review or interpretation. However, each patient chart was manually reviewed by a single neurologist, so we are unable to assess the reliability or misclassification bias of individual abstractors.

CONCLUSIONS

Individual medical history elements in the GWTG-Stroke registry were only moderately reliable in this single-center study, suggesting that caution should be exercised when relying heavily on any individual history elements in registry-based research. Combining these variables to calculate the CHA_2DS_2-VASc score was somewhat more reliable. Using GWTG-Stroke registry data to estimate the CHA_2DS_2-VASc score may be particularly reasonable because the score is often used to estimate risk category (ie, low versus high risk), in which case some degree of inaccuracy may be acceptable. Multicenter data are needed before assuming generalizability.

ARTICLE INFORMATION

Received January 11, 2022; accepted April 8, 2022.

Affiliation

Department of Neurology, University of Pennsylvania, Philadelphia, PA.

Sources of Funding

This work was supported by the AHA (19CDA-34630033, C.G.F.) and the National Institutes of Health (K23-NS110993, C.G.F.).

Disclosures

None.

Supplemental Material

Data S1
Figure S1

REFERENCES

1. Schwamm LH, Fonarow GC, Reeves MJ, Pan W, Frankel MR, Smith EE, Ellrodt G, Cannon CP, Liang LI, Peterson E, et al. Get With the Guidelines-Stroke is associated with sustained improvement in care for patients hospitalized with acute stroke or transient ischemic attack. *Circulation*. 2009;119:107–115. doi: 10.1161/CIRCULATIONAHA.108.783688
2. Fonarow GC, Reeves MJ, Smith EE, Saver JL, Zhao X, Olson DW, Hernandez AF, Peterson ED, Schwamm LH; Committee GW-SS. Characteristics, performance measures, and in-hospital outcomes of the first one million stroke and transient ischemic attack admissions in get with the guidelines-stroke. *Circ Cardiovasc Qual Outcomes*. 2010;3:291–302. doi: 10.1161/CIRCOUTCOMES.109.921858
3. Ormseth CH, Sheth KN, Saver JL, Fonarow GC, Schwamm LH. The American Heart Association's Get With the Guidelines (GWTG)-Stroke development and impact on stroke care. *Stroke Vasc Neurol*. 2017;2:94–105. doi: 10.1136/svn-2017-000092
4. Xian Y, Fonarow GC, Reeves MJ, Webb LE, Blevins J, Demyanenko VS, Zhao X, Olson DM, Hernandez AF, Peterson ED, et al. Data quality in the American Heart Association Get With The Guidelines-Stroke (GWTG-Stroke): results from a national data validation audit. *Am Heart J*. 2012;163:392–398, 398.e391. doi: 10.1016/j.ahj.2011.12.012
5. Ehrlich ME, Liang LI, Xu H, Kosinski AS, Hernandez AF, Schwamm LH, Smith EE, Fonarow GC, Bhatt DL, Peterson ED, et al. Intravenous tissue-type plasminogen activator in acute ischemic stroke patients with history of stroke plus diabetes mellitus. *Stroke*. 2019;50:1497–1503. doi: 10.1161/STROKEAHA.118.024172
6. Kaufman BG, Shah S, Hellkamp AS, Lytle BL, Fonarow GC, Schwamm LH, Lesén E, Hedberg J, Tank A, Fita E, et al. Disease burden following non-cardioembolic minor ischemic stroke or high-risk TIA: a GWTG-Stroke Study. *J Stroke Cerebrovasc Dis*. 2020;29:105399. doi: 10.1016/j.jstrokecerebrovasdis.2020.105399
7. Reeves MJ, Vaidya RS, Fonarow GC, Liang L, Smith EE, Matulonis R, Olson DM, Schwamm LH; Get With The Guidelines Steering C, Hospitals. Quality of care and outcomes in patients with diabetes hospitalized with ischemic stroke: findings from Get With the Guidelines-Stroke. *Stroke*. 2010;41:e409–e417. doi: 10.1161/STROKEAHA.109.572693
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. doi: 10.1161/CIRCULATIONAHA.109.932822
9. Olesen JB, Torp-Pedersen C, Hansen ML, Lip GY. The value of the CHA_2DS_2-VASc score for refining stroke risk stratification in patients with atrial fibrillation with a CHADS2 score 0–1: a nationwide cohort study. *Thromb Haemost*. 2012;107:1172–1179. doi: 10.1160/TH12-03-0175
10. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the Euro Heart Survey on Atrial Fibrillation. *Chest*. 2010;137:263–272. doi: 10.1378/chest.09-1584
11. Lewis WR, Fonarow GC, Grau-Sepulveda MV, Smith EE, Bhatt DL, Hernandez AF, Olson D, Peterson ED, Schwamm LH. Improvement in use of anticoagulation therapy in patients with ischemic stroke: results from Get With The Guidelines-Stroke. *Am Heart J*. 2011;162:692–699.e692. doi: 10.1016/j.ahj.2011.07.019

SUPPLEMENTAL MATERIAL

Data S1. Methods - Instructions for Abstractors

Evaluate each patient for the presence (score as a 1) or absence (score as a 0) of the following previously known medical history elements:

1. Diabetes Mellitus
2. Hypertension
3. Heart Failure
4. Previous Stroke
5. Previous TIA
6. Vascular Disease

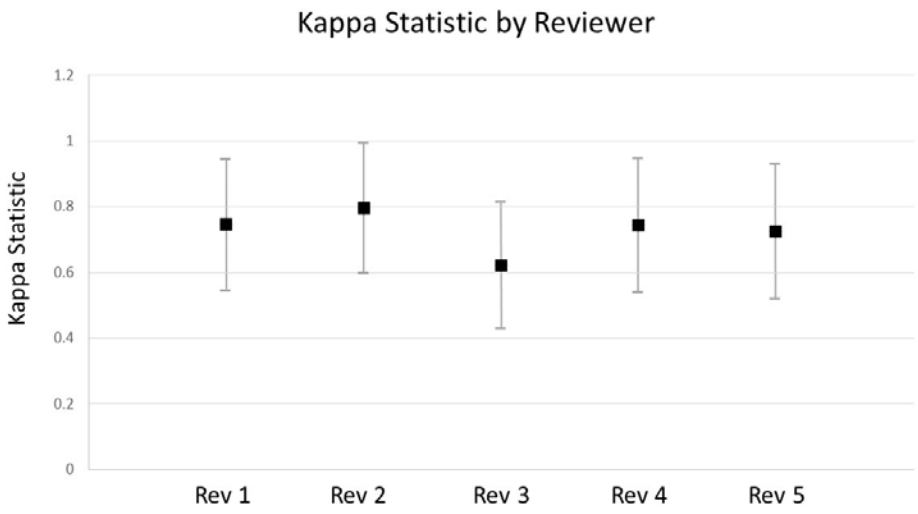
Previously known elements of the medical history are defined as conditions known to exist prior to the admission for the acute stroke. Conditions that are newly diagnosed during the stroke hospitalization are not considered part of the previously known medical history and should instead be categorized as newly diagnosed (separate variables in the dataset). The medical history elements of interest are described in more detail below:

1. **Diabetes Mellitus** – Based on a history of physician diagnosed diabetes mellitus, regardless of duration of disease or treatment. Do not score this element based on laboratory data alone.
2. **Hypertension** – Based on documented history of high blood pressure regardless of treatment status. Patients who are on anti-hypertensive medications prior to admission should be considered to have a history of hypertension. Do not consider blood pressure readings during the stroke admission.
3. **Heart failure** – Based on a documented history of heart failure known prior to the index admission. This includes heart failure with either reduced or preserved ejection fraction. This is not solely based on echocardiographic data, but requires a clinical diagnosis prior to index stroke admission.
4. **Previous Stroke** (**or TIA**)– Based on a documented history of stroke (or TIA) prior to the index hospitalization. The index stroke admission does not constitute a history of stroke. If the patient was first admitted to another facility and subsequently transferred to our facility, that does not constitute a previous history of stroke.
5. **Vascular Disease** – Based on a documented history of previous myocardial infarction, peripheral arterial disease, or aortic plaque.

In assessing medical history, review materials from the index admission, including:

Admission notes, consultation notes, transfer notes, any *ED documentation,* including physician notes, nursing notes, ED pathway documentation, ED triage notes, ED registration, ED flowsheets, and any *inpatient documentation*, including H&P, daily progress notes, consultation notes, medication reconciliation, nursing progress notes, nursing assessments, multidisciplinary notes (i.e. PT/OT, speech therapy, nutrition team).

Figure S1. Reliability of CHA_2DS_2-VASc score across five reviewers



Reliability of CHA_2DS_2-VASc score across five reviewers: Reliability of CHA_2DS_2-VASc scores is assessed by quadratic weight kappa. Five physicians reviewed between 90 and 95 unique charts. Kappa statistics and 95% confidence intervals were calculated for each reviewer. Reliability was similar across reviewers.