J Neurosurg 121:482–489, 2014

Analysis of subarachnoid hemorrhage using the Nationwide Inpatient Sample: the NIS-SAH Severity Score and Outcome Measure

Clinical article

Chad W. Washington, M.D., M.S., M.P.H.S.,[1,3] Colin P. Derdeyn, M.D.,[2,3] Ralph G. Dacey Jr., M.D.,[1] Rajat Dhar, M.D.,[3] and Gregory J. Zipfel, M.D.[1,3]

Departments of [1]Neurological Surgery, [2]Radiology, and [3]Neurology, Washington University in St. Louis, Missouri

Object. Studies using the Nationwide Inpatient Sample (NIS), a large ICD-9–based (International Classification of Diseases, Ninth Revision) administrative database, to analyze aneurysmal subarachnoid hemorrhage (SAH) have been limited by an inability to control for SAH severity and the use of unverified outcome measures. To address these limitations, the authors developed and validated a surrogate marker for SAH severity, the NIS-SAH Severity Score (NIS-SSS; akin to Hunt and Hess [HH] grade), and a dichotomous measure of SAH outcome, the NIS-SAH Outcome Measure (NIS-SOM; akin to modified Rankin Scale [mRS] score).

Methods. Three separate and distinct patient cohorts were used to define and then validate the NIS-SSS and NIS-SOM. A cohort (n = 148,958, the "model population") derived from the 1998–2009 NIS was used for developing the NIS-SSS and NIS-SOM models. Diagnoses most likely reflective of SAH severity were entered into a regression model predicting poor outcome; model coefficients of significant factors were used to generate the NIS-SSS. Nationwide Inpatient Sample codes most likely to reflect a poor outcome (for example, discharge disposition, tracheostomy) were used to create the NIS-SOM.

Data from 716 patients with SAH (the "validation population") treated at the authors' institution were used to validate the NIS-SSS and NIS-SOM against HH grade and mRS score, respectively.

Lastly, 147,395 patients (the "assessment population") from the 1998–2009 NIS, independent of the model population, were used to assess performance of the NIS-SSS in predicting outcome. The ability of the NIS-SSS to predict outcome was compared with other common measures of disease severity (All Patient Refined Diagnosis Related Group [APR-DRG], All Payer Severity-adjusted DRG [APS-DRG], and DRG).

Results. The NIS-SSS significantly correlated with HH grade, and there was no statistical difference between the abilities of the NIS-SSS and HH grade to predict mRS-based outcomes. As compared with the APR-DRG, APS-DRG, and DRG, the NIS-SSS was more accurate in predicting SAH outcome (area under the curve [AUC] = 0.69, 0.71, 0.71, and 0.79, respectively).

A strong correlation between NIS-SOM and mRS was found, with an agreement and kappa statistic of 85% and 0.63, respectively, when poor outcome was defined by an mRS score > 2 and 95% and 0.84 when poor outcome was defined by an mRS score > 3.

Conclusions. Data in this study indicate that in the analysis of NIS data sets, the NIS-SSS is a valid measure of SAH severity that outperforms previous measures of disease severity and that the NIS-SOM is a valid measure of SAH outcome. It is critically important that outcomes research in SAH using administrative data sets incorporate the NIS-SSS and NIS-SOM to adjust for neurology-specific disease severity.
(http://thejns.org/doi/abs/10.3171/2014.4.JNS131100)

Key Words • **subarachnoid hemorrhage** • **aneurysm** • **skull base** • **Nationwide Inpatient Sample** • **Hunt and Hess grade** • **vascular disorders** • **modified Rankin Scale score**

Abbreviations used in this paper: AHRQ = Agency for Healthcare Research and Quality; APR-DRG = All Patient Refined Diagnosis Related Group; APS-DRG = All Payer Severity-adjusted DRG; AUC = area under the curve; CCI = Charlson Comorbidity Index; HH = Hunt and Hess; ICD-9-CM = International Classification of Diseases, Ninth Revision, Clinical Modification; mRS = modified Rankin Scale; NIS = Nationwide Inpatient Sample; NIH-SOM = NIS-SAH Outcome Measure; NIH-SSS = NIH-SAH Severity Score; ROC = receiver operating characteristic; SAH = subarachnoid hemorrhage; WFNS = World Federation of Neurosurgical Societies.

The Nationwide Inpatient Sample (NIS) is one of the databases developed as part of the Healthcare Cost and Utilization Project (HCUP) in an effort to provide accurate patient-level data to improve the quality and effectiveness of health care (http://www.hcup-us.ahrq.gov/overview.jsp). Data contained in the NIS includes diagnoses and procedures (using ICD-9-CM [International Classification of Diseases, Ninth Revision, Clinical Modification] codes), patient demographics, to-

tal charges, length of stay, discharge status, and hospital characteristics. The NIS collects information on hospital stays from more than 1000 hospitals, representing 20% of all United States community-based inpatient health care facilities (http://www.hcup-us.ahrq.gov/nisoverview.jsp). These data hold great potential for evaluating variables that impact health care outcomes across broad populations.

Aneurysmal subarachnoid hemorrhage (SAH) is a heterogeneous disease complicated by poor outcomes and is uncommon enough that adequately sized studies require large multicenter cohorts. To date, more than 20 studies have used data from the NIS to examine a variety of important questions regarding SAH.[3,6,8,9,12–14,20,23–25,29–31,34,35,37,42–44,46,51,53] However, all of these analyses suffer from a critical limitation: neurological status was not controlled for when comparing outcomes across patients and cohorts. This limitation results from the absence of direct measures of SAH severity in the NIS data and can be significant given that SAH severity (for example, Hunt and Hess [HH] grade[26]) is one of the most important determinants of patient outcome after SAH.[26,47,58] Instead, investigators have used nonspecific indicators of comorbidity and disease severity to adjust patient analyses.[3,6,8,9,12–14,20,23–25,29–31,34,35,37,42–44,46,51,53] Another significant limitation of past NIS-based studies of SAH is that the applied outcome measures have not been validated for this condition and have relied solely on the discharge disposition of the patient, with discharge home being a "good" outcome and all other discharges being "poor." These critical methodological limitations have very likely curbed the clinical relevance of conclusions drawn from the NIS-based SAH studies. While the NIS supplies data for large-scale SAH research, it is imperative that disease-specific severity adjustment and outcome measures are developed to permit appropriate statistical analysis and meaningful clinical interpretation.

The hypothesis of the present study was that information exists within the NIS database that allows for the calculation of an SAH severity score (akin to the HH grade) and an SAH outcome measure (akin to the modified Rankin Scale [mRS] score[28]). Specifically, our objectives were to 1) develop the NIS-SAH Severity Score (NIS-SSS) and the NIS-SAH Outcome Measure (NIS-SOM) based on data available in the NIS, 2) validate the NIS-SSS and NIS-SOM against established SAH grading systems (HH classification and mRS, respectively) based on data from a cohort of SAH patients from our institution, and 3) compare the performance of the NIS-SSS and NIS-SOM against previously used methods for controlling for disease severity in the NIS.

Methods

Patient Population

To ensure the highest statistical fidelity possible, three separate patient cohorts were used to define and then validate the NIS-SOM and the NIS-SSS. A separate population was used for each portion of the study: 1) model creation, 2) model validation, and 3) performance assessment. An initial cohort derived from the 1998–2009 NIS was used for developing the NIS-SSS and NIS-SOM models, termed the "model population" (n = 148,958 patients). For model validation, data from SAH patients treated at our institution (for whom HH grades and mRS scores were available) were used, composing the "validation population" (n = 716 patients). Lastly, unique patients from the 1998–2009 NIS, independent of those in the model population, were used for assessing the performance of the NIS-SSS (the "assessment population," n = 147,395 patients).

The model and assessment population were derived from the 1998–2009 NIS databases. Patients were selected if their primary diagnosis was aneurysmal SAH (ICD-9-CM 430). Patients were excluded if they had diagnoses consistent with head trauma (ICD-9-CM 800.0–801.9, 803.0–804.9, 850.0–854.1, 873.0–873.9) or arteriovenous malformation and/or fistula (ICD-9-CM diagnosis 747.81; ICD-9-CM procedures 39.53, 92.30). Patients were also excluded if their length of stay was less than 1 day with an associated discharge to home. The final cohort was randomly divided into the model population (n = 148,958) and the assessment population (n = 147,395) using the RANUNI function provided by SAS version 9.3 (SAS Institute Inc.).

The validation population (n = 716) was derived from a database at the authors' institution containing the billing data (ICD-9-CM diagnosis and procedure codes) for all hospitalizations from January 2002 through September 2011. Patients were selected based on the admitting diagnosis of aneurysmal SAH (ICD-9-CM 430). The charts of these patients were reviewed to verify that the reason for admission was a ruptured cerebral aneurysm. Those patients without an aneurysm or with other etiologies (for example, trauma, arteriovenous malformation, hemorrhagic stroke) were excluded. The final population was then cross-matched to their admission HH grade and hospital discharge mRS score, which were obtained from the prospectively collected database maintained as part of our institutional Stroke and Cerebrovascular Center.

Previous Methods of Case-Mix Adjustment

The 4 most standard methods for case-mix adjustment in administrative data are the Charlson Comorbidity Index (CCI),[11] Agency for Healthcare Research and Quality (AHRQ) comorbidity measures,[16] All Patient Refined Diagnosis Related Groups (APR-DRGs) severity and mortality indices (3M Health Information Systems),[4] and All Payer Severity-adjusted Diagnosis Related Groups (APS-DRGs) mortality weight (HSS Inc.).[2] The CCI is a summed score calculated from the presence of comorbid conditions and is intended to predict patient mortality. The AHRQ comorbidity measures assign variables for a number of predefined comorbidities (that is, hypertension, diabetes, renal disease, and so forth). The APR-DRGs and the APS-DRGs are modifications to the Medicare/Medicaid DRG prospective payment system. Both use similar information regarding patient comorbidities to calculate an index or weight to account for a patient's severity of illness. The methodology used for their calculations have been described,[2,4] but the software is proprietary. None

of these techniques uses neurological condition in their calculation.

Definition and Validation: NIS-SSS

The NIS-SSS was developed using both the HH[26] and the World Federation of Neurosurgical Societies (WFNS)[55] SAH grading scales as templates. With these SAH-specific measures as guidelines, the model population was interrogated for the ICD-9-CM diagnosis codes for coma (780.01, 780.03), stupor (780.02, 780.09), hydrocephalus (331.3, 331.4), paresis/plegia (438.2–438.53, 781.4), aphasia (438.1–438.89), and cranial nerve deficits (378.5–378.56, 379.4–379.43). Patients were also categorized based on ICD-9-CM procedure codes for the treatment of hydrocephalus (ventriculostomy 02.2 and CSF shunting 02.31–02.39) and mechanical ventilation (96.04, 96.7–96.72). We assumed that requiring mechanical ventilation and/or having hydrocephalus would serve as surrogates for decreased levels of consciousness.

The ability of each of the above categories to predict a poor outcome was then evaluated using chi-square analysis. Those diagnoses found to be significant univariate predictors were then entered into a multivariate logistic regression model. This generated coefficients (B_i) for significant predictors, which were used to obtain the NIS-SSS for each patient using the following equation:

$$e^{\left(\sum_{i=1}^{7} B_i * x_i\right)},$$

where $x_i = 1$ when the diagnosis is present and 0 if absent. This equation was derived empirically by comparing multiple formulations.

The concurrent validity of this NIS-SSS model was assessed within the validation population by analyzing how well it matched HH grade using 1-way ANOVA and ordinal logistic regression. The abilities of both the NIS-SSS and HH grading to predict mRS score at patient discharge were compared using logistic regression.

NIS-SSS Compared With Previous Measures

The NIS-SSS was calculated for each patient in the assessment population, and the abilities of the NIS-SSS, APR-DRG mortality index, APR-DRG severity index, APS-DRG mortality weight, and DRG to predict NIS-SOM were compared. The evaluation was accomplished by comparing the receiver operating characteristic (ROC) curve association statistics from a logistic regression model. The CCI and AHRQ comorbidity measures were purposely excluded from these comparisons; their calculations are clearly related to specific comorbid conditions and by their definition have no relation to neurological condition.

Definition and Validation: NIS-SOM

A dichotomous outcome variable, the NIS-SOM, was created. Good outcome was defined as discharge to home or a rehabilitation facility and/or hospital. Poor outcome was defined as in-hospital mortality; discharge to a nursing facility, extended care facility, or hospice; placement of a tracheostomy tube (ICD-9-CM procedures: 31.1, 31.2. 31.21, 31.29), and/or placement of a gastrostomy tube (ICD-9-CM procedures: 43.1, 43.11, 43.19, 44.32, 44.38, 44.39). This definition expands on previously used NIS outcome measures, which consider only discharge status, by including tracheostomy and gastrostomy tube placement as additional indications of disability. This definition of outcome was then applied to each patient in the validation population (for whom mRS score at discharge was available). Agreement between the NIS-SOM and mRS score was then assessed using the kappa statistic (with both mRS scores > 2 and mRS scores > 3 considered as poor outcomes).[33] This analysis was also completed using the previously defined discharge-based outcome, allowing us to assess the performance of the new NIS-SOM.

Statistical Analysis

All statistical analyses were completed using SAS version 9.3 (SAS Institute Inc.). Statistical significance was defined as $p < 0.01$. For comparisons of interrater reliability using the kappa statistic, $\kappa \geq 0.61$ was considered substantial and $\kappa \geq 0.81$ was considered almost perfect.[33]

Results

Development and Validation: NIS-SSS

Factors found to be significant predictors of NIS-SOM in the model population are shown in Table 1. These factors were entered into the multivariate logistic regression model and were all found to be significant predictors of outcome and formed the basis of the NIS-SSS.

Using the validation population, NIS-SSS strongly correlated with HH grade (Fig. 1). According to regression analysis, concordance was 67.5%, discordance was 17.5%, and the area under the curve (AUC) was 0.75. The difference in the mean NIS-SSS was statistically significant across all HH grades except between Grades I and II.

There was no significant difference between NIS-SSS and HH grade in predicting poor outcome as defined by the mRS score at discharge (Fig. 2). The AUC for NIS-SSS was 0.81 (mRS score > 2) and 0.89 (mRS score > 3). The AUC for HH grade was 0.79 and 0.84 for mRS score > 2 and mRS score > 3, respectively (Table 2).

Comparing NIS-SSS With APR-DRG, APS-DRG, and DRG Systems

Results from the assessment population (AUC for each measure follows) comparing NIS-SSS (0.79) with APR-DRG mortality index (0.75), APR-DRG severity index (0.69), APS-DRG mortality weight (0.70), and DRG (0.71) demonstrated that NIS-SSS is significantly more accurate in predicting outcome when comparing AUC, rates of concordance, and rates of discordance (Fig. 3 and Table 3).

NIS-SOM Validation

In the validation population, the agreement between poor outcome as defined by an mRS score > 2 and NIS-SOM was 85.2% with a kappa statistic of 0.63 (substantial agreement; that is, $\kappa \geq 0.61$). The agreement between

TABLE 1: Results from univariate and multivariate analyses of factors impacting SAH outcomes*

Factor	Univariate		Multivariate		
	OR	95% CI	OR	95% CI	p Value
mechanical ventilation	8.57	8.31–8.34	7.56	7.37–7.53	<0.0001
hydrocephalus	1.46	1.41–1.51	0.93	0.90–0.96	<0.0001
treatment of hydrocephalus	1.79	1.74–1.84	1.25	1.21–1.29	<0.0001
coma	8.28	7.77–8.82	6.01	5.65–6.41	<0.0001
stupor	0.89	0.76–1.03	NA	NA	0.13
cranial nerve palsy	0.32	0.28–0.36	0.32	0.29–0.36	<0.0001
paralysis/paraparesis	1.49	1.34–1.66	1.66	1.46–1.88	<0.0001
aphasia	1.31	1.17–1.48	1.56	1.36–1.79	<0.0001

* NA = not applicable.

NIS-SOM and poor outcome as defined by an mRS score > 3 was 94.8% with a kappa statistic of 0.84 (almost perfect agreement; that is, $\kappa \geq 0.81$). The distribution of each mRS category to "good" and "poor" NIS-SOM is shown in Table 4.

Agreement between previously used purely discharge-based outcome measures and poor outcome as defined by an mRS score > 2 was 80% with a kappa statistic of 0.47 (moderate agreement; that is, $\kappa \geq 0.41$). Agreement between discharge-based outcome and poor outcome as defined by an mRS score > 3 was 91% with a kappa statistic of 0.70 (substantial agreement, $\kappa \geq 0.61$).

Discussion

To meaningfully compare outcomes across large cohorts in a heterogeneous disease like SAH, it is imperative to have disease-specific severity measures to adjust for case mix; it is also critical to have outcome measures with validated clinical relevance to disability. The NIS-SSS and NIS-SOM provide new means of optimizing SAH research when utilizing data contained within the NIS, which otherwise lacks SAH-specific measures like HH grade and mRS score.

We derived and validated an SAH-specific severity score, the NIS-SSS, which can be calculated from information provided in administrative databases. Our results demonstrated a strong correlation between NIS-SSS and HH grade, one of the most commonly used measures of SAH severity. Moreover, there was no statistical difference between NIS-SSS and HH grade in predicting mRS-based outcome, confirming the NIS-SSS's utility in clinical research using administrative databases. Furthermore, as compared with previous NIS measures of severity, which are based on patient comorbidities and not neurological condition, the NIS-SSS was shown to be a statistically superior predictor of post-SAH outcomes. From previous studies we know that initial presentation is one of the most significant predictors of outcome and a critical variable to adjust for in an analysis.[47] Until now, research using the NIS to study SAH has suffered from an inability to control for severity. The NIS-SSS represents a significant improvement in accounting for factors related to SAH severity in database research.

Along with the NIS-SSS, we created an improved measure for assessing SAH outcomes in the NIS, that is, the NIS-SOM. The mRS is an established outcome measure in stroke research; outcome is usually dichotomized into "good" (mRS score ≤ 2) versus "poor" (mRS score ≥ 3), as was done in the International Subarachnoid An-

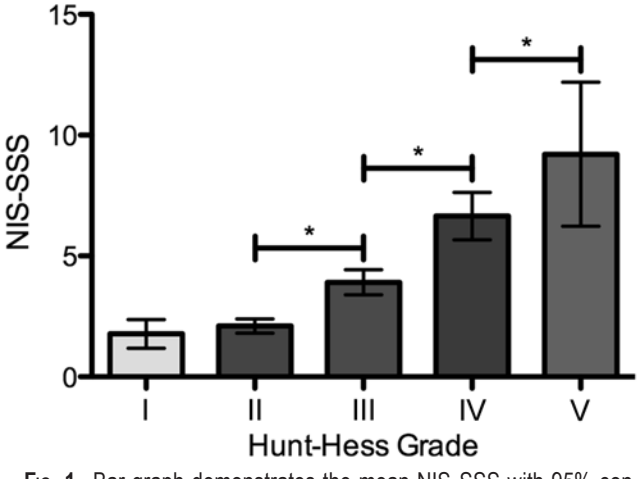

Fig. 1. Bar graph demonstrates the mean NIS-SSS with 95% confidence interval *error bars* for each HH grade. The difference in mean NIS-SSS was statistically significant across HH Grade II to III, Grade III to IV, and Grade IV to V. There was no statistical difference between Grade I and II. *Asterisks* indicate statistically significant difference ($p < 0.005$, 1-way ANOVA).

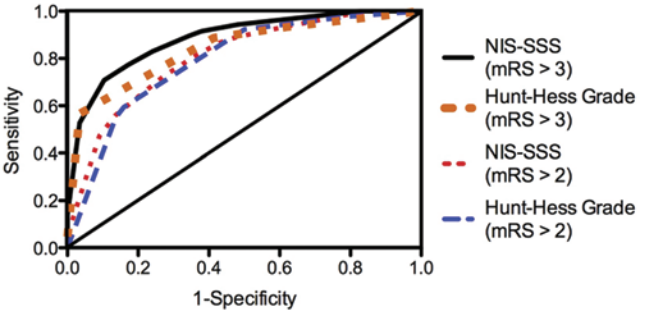

Fig. 2. The mRS score was used to define "poor" outcome. In the first model, poor outcome was defined as mRS score > 2; in the second model, poor outcome was defined as mRS score > 3. *Curves* represent the ROC curve profiles for NIS-SSS and HH grade to predict these outcomes.

TABLE 2: Comparison of NIS-SSS and HH grade in predicting outcomes

Measure	ROC AUC	Concordance	Discordance	ROC Difference*	p Value for ROC Difference
poor outcome, mRS score >2					
NIS-SSS	0.806	75.6	14.5	−0.0135	0.7623
HH	0.792	68.8	10.3		
poor outcome, mRS score >3					
NIS-SSS	0.885	85.3	8.3	−0.0434	0.3858
HH	0.841	75.4	7.2		

* ROC difference = HH − NIS-SSS.

eurysm Trial (ISAT)—but outcome as expressed by the mRS score is lacking in most administrative databases, hampering evaluation of meaningful outcomes.[41] Using measures available in the NIS, we have shown that the NIS-SOM has substantial agreement with dichotomized mRS outcomes in our institutional data set (85.2% agreement and $\kappa = 0.63$). Because the interrater reliability for mRS score can vary significantly ($\kappa = 0.25$–0.72)[45]—particularly at the interface between mRS scores of 2 and 3—we also examined agreement between our NIS-SOM and "poor" outcome as defined by an mRS score ≥ 4. In this analysis, the kappa value demonstrated near perfect agreement with $\kappa = 0.84$ and 94.8% agreement. Further, by expanding the definition of poor outcome to include tracheostomy and/or gastrostomy tube placement, agreement with mRS-based outcomes was improved as compared with the previously used discharge-based outcome measure. These results indicate that the NIS-SOM is a valid and improved measure for assessing patient outcome after SAH when using the NIS.

The NIS database has been extensively studied in regard to SAH and aneurysm treatment.[3,6,8,9,12–14,20,23–25,29–31,34,35,37,42–44,46,51,53] These studies have assessed the impact of hospital characteristics,[3,13,14,29,34,43,53] mode of treatment,[23,25,51] patient comorbidities,[20,24,35] and patient demographics.[6,8,32,33,45] For example, a number of studies have found that hospital characteristics, such as treatment volume,[13,34,43] academic status,[3] and the ability to provide angioplasty,[29] significantly affect outcomes. Others have shown that surgical clipping is associated with a higher use of hospital resources,[23,51] the increased incidence of postprocedure epilepsy,[25] and increased mortality.[51] An important factor potentially confounding these results is that, thus far, case-mix adjustment has involved the use of the CCI,[8,9,13,35,42,44] AHRQ comorbidity measures,[6,12,14,20,23–25,31,46,53] APR-DRGs,[29,43,51] or nothing at all.[3,30,34,37] The CCI and AHRQ comorbidity measures were developed to control for the impact of patient comorbidities on outcomes. After using the APR-DRG index, some authors stated that direct comparisons between the APR-DRG and HH grade can be made.[43] While others have shown that APR-DRG does accurately predict in-hospital mortality[52] and outperforms the Medicare and Medicaid DRG system in predicting resource utilization,[19] its direct correlation with HH grade has not been demonstrated. All of the aforementioned methodologies are based on patient-specific factors unrelated to neurological condition, such as patient age, comorbidities (hypertension, diabetes, coronary artery disease, and so forth), and procedures. It has been shown that these factors are important in determining patient outcomes; however, none account for the severity of SAH. The NIS-SSS resolves this issue by evaluating clinical factors more specific to patients with SAH.

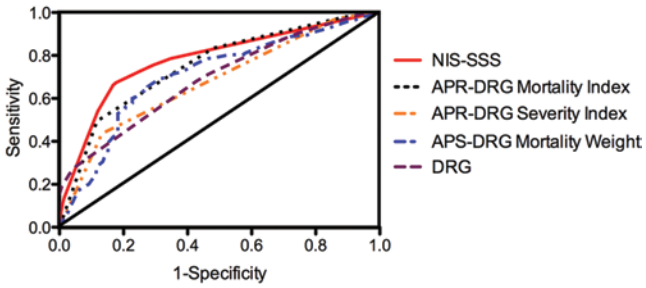

Fig. 3. The ROC curve profiles of the NIS-SSS, APR-DRG mortality index, APR-DRG severity index, APS-DRG mortality weight, and DRG to predict NIS-SOM. The NIS-SSS outperformed all of the standard measures provided through the NIS.

Beyond clinical research, significant focus is being placed on measuring health care outcomes and quality. In 1983, the DRG prospective payment system (PPS) was implemented as the nationwide reimbursement system for Medicare and Medicaid.[36,38] Soon after its introduction, the DRG's use was expanded to measure health care outcomes.[56] Numerous opponents argue that DRGs are inadequate in capturing a true case mix and that more accurate methods of assessing disease-specific factors must be developed.[15,17,22,27,39,49,50,54] This opposition led to the development of a number of severity measures like the APR-DRG.

The APR-DRG is emerging as the case-mix adjustment technique of choice in the analysis of administrative data, as 30 states and the District of Columbia have adopted its use for risk adjustment in determining reimbursement and/or measurement of quality (http://solutions.3m.com/wps/portal/3M/en_US/Health-Information-Systems/HIS/Our-Partners/State-Initiatives/).[5] The AHRQ, a branch of the US Department of Health and Human Services, developed another method of case-mix adjustment. They have proposed using inpatient mortality, risk adjusted by the APR-DRG method, as a measure of clinical quality in craniotomy.[1] With the growing socioeconomic interest in health care quality and cost-effectiveness, measures such as these are being implemented with increasing frequen-

TABLE 3: Receiver operating characteristic curve association statistics for predicting poor outcome in NIS assessment population

Measure	ROC AUC	Concordance	Discordance	Difference in ROC	p Value for ROC Difference
NIS-SSS	0.7873	69.6	13.4	—	—
APR-DRG mortality index	0.7508	62.9	12.8	−0.0365	<0.001
APR-DRG severity index	0.6872	45.2	9.6	−0.1090	<0.001
APS-DRG mortality weight	0.7059	69.3	28.1	−0.0814	<0.001
DRG	0.7111	59.3	21.1	−0.0761	<0.001

cy to develop clinical guidelines and quality tools and as a basis for reimbursement.[10] Considering the serious implications, we must ensure that these measures accurately represent important clinical factors established through years of research and observation. To this point, as the use of administrative data in health care research continues to expand,[18,32,40] validated methods to control for presenting neurological conditions in other cerebrovascular diseases, such as ischemic stroke and intracerebral hemorrhage, should be developed.

The NIS-SSS is unique in that it considers features more specific to SAH and neurological outcomes. However, there are significant limitations to this method, which are inherent to the data used to develop it. First, our study was retrospective and used a very heterogeneous population of patients. Therefore, we made every attempt to account for these facts by comparing our measures, both the NIS-SSS and NIS-SOM, with known standards, HH grade and mRS score, in a population with prospectively collected data. Moreover, creating the model in one patient population and then validating it in two separate cohorts helped to ensure that we did not over-fit our model. Second, diagnoses and outcomes were calculated based on ICD-9-CM codes acquired from billing data. Billing data are generated following patient discharge. We realize that it is possible that some of the diagnoses and procedures occurred after presentation; for example, it is possible that a patient was admitted in good neurological condition but that their poor outcome was related to an iatrogenic complication. Similarly, a patient presenting in a poor neurological condition may have ultimately had a good outcome after successful treatment. Unfortunately, this cannot be controlled for within the NIS, where the temporal component of data is unknown. The validation population, with its prospectively collected HH grades and mRS outcome scores, was used to account for the unknown temporal component. The significant, although not perfect, correlation between HH grade and NIS-SSS and between mRS score and NIS-SOM provides evidence that this limitation does not negate the utility of these new measures. Nonetheless, the lack of full temporal knowledge may continue to limit the complete application of administrative data research.

TABLE 4: Breakdown of validation population according to outcome classifications per the NIS-SOM for each mRS category

	NIS-SOM	
mRS Score	Good Outcome (%)	Poor Outcome (%)
0	100	0
1	96	4
2	100	0
3	93	7
4	44	56
5	0	100
6	0	100

Further limitations include the inherent errors that exist within administrative data; for instance, it is very likely that not all patients from the NIS were truly admitted for aneurysmal SAH. Fortunately, research using administrative data has shown that the inaccuracy related to ICD-9-CM coding can be significantly improved through the use of multiple related codes,[21,57] which was our methodology. Lastly, outcomes were based on factors at the time of patient discharge rather than a more expanded follow-up period. While this strategy is imperfect, previous administrative studies have provided justification by demonstrating that a facility's in-hospital mortality calculated from administrative data are an accurate surrogate for 30-day mortality.[7,48]

It is important for the reader to appreciate the context in which the NIS-SSS and NIS-SOM were created. They were not developed to replace HH grade, WFNS grade, or mRS score, which are important and validated clinical tools. Further, many additional factors unknown in administrative data—that is, imaging findings, treatment complications, development of cerebral vasospasm—ultimately contribute to outcome following SAH. The goal of this study was to develop a methodology through which research in SAH using administrative data could be made more accurate and clinically applicable than the methods used currently.

Conclusions

We developed and validated the NIS-SSS and NIS-SOM, two original measures that allow controlling for the severity of SAH and accurately determining outcome when utilizing the NIS in clinical research. The NIS-SSS outperformed previously established measures of disease severity in predicting outcome. The NIS-SSS holds promise as a superior means of adjusting for case mix when comparing surgical versus endovascular treatments as well as outcomes related to hospital characteristics. Future stud-

ies analyzing the NIS-SSS's ability to predict costs and resource utilization will be initiated, as this score may have a significant impact on measurements of quality and potential reimbursement. We believe that the NIS-SSS and NIS-SOM should be included in all future NIS analyses.

Disclosure

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper. Dr. Derdyn is a consultant for W. L. Gore & Associates, holds stock in Pulse Therapeutics, and has received non–study-related support from MicroVention Inc. and Penumbra Inc. Research reported in this publication was supported by the Washington University Institute of Clinical and Translational Sciences Grant No. UL1 TR000448 from the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH). The content is solely the responsibility of the authors and does not necessarily represent the official view of the NIH.

Author contributions to the study and manuscript preparation include the following. Conception and design: Washington, Derdeyn. Acquisition of data: Washington. Analysis and interpretation of data: Washington, Dhar, Zipfel. Drafting the article: Washington. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Washington. Statistical analysis: Washington. Study supervision: Zipfel.

References

1. Agency for Healthcare Research and Quality: **Craniotomy Mortality Rate. Inpatient Quality Indicators #13. Technical Specifications. Provider-Level Indicator. Procedure Mortality Indicator. AHRQ Quality Indicators, Version 4.3, August 2011.** Rockville, MD: AHRQ, 2011
2. **All Payer Severity-Adjusted DRGs (APS-DRGs®): Normalized Charge, LOS, and Mortality Weights. Version 0402H (for APS-DRGs Version 21.0).** HSS, Inc, 2004 (http://www.hcup-us.ahrq.gov/db/nation/nis/APS-DRGs%20V21%20Weights%20Manual.pdf) [Accessed May 1, 2014]
3. Andaluz N, Zuccarello M: Recent trends in the treatment of cerebral aneurysms: analysis of a nationwide inpatient database. **J Neurosurg 108:**1163–1169, 2008
4. Averill RF, Goldfield N, Hughes JS, Bonazelli J, McCullough EC, Steinbeck BA, et al: **All Patient Refined Diagnosis Related Groups (APR-DRGs), Version 20.0. Methodology Overview.** Wallingford, CT: 3M Health Systems, 2003 (http://www.hcup-us.ahrq.gov/db/nation/nis/APR-DRGsV20MethodologyOverviewandBibliography.pdf) [Accessed May 1, 2014]
5. Bachrach D: Payment reform: creating a sustainable future for Medicaid. Policy brief. **Center for Health Care Strategies, Inc.** (http://www.chcs.org/usr_doc/Medicaid_Payment_Reform_Brief.pdf) [Accessed May 1, 2014]
6. Bateman BT, Olbrecht VA, Berman MF, Minehart RD, Schwamm LH, Leffert LR: Peripartum subarachnoid hemorrhage: nationwide data and institutional experience. **Anesthesiology 116:**324–333, 2012
7. Borzecki AM, Christiansen CL, Chew P, Loveland S, Rosen AK: Comparison of in-hospital versus 30-day mortality assessments for selected medical conditions. **Med Care 48:**1117–1121, 2010
8. Brinjikji W, Lanzino G, Rabinstein AA, Kallmes DF, Cloft HJ: Age-related trends in the treatment and outcomes of ruptured cerebral aneurysms: a study of the nationwide inpatient sample 2001-2009. **AJNR Am J Neuroradiol 34:**1022–1027, 2013
9. Brinjikji W, Rabinstein AA, Lanzino G, Cloft HJ: Racial and ethnic disparities in the treatment of unruptured intracranial aneurysms: a study of the Nationwide Inpatient Sample 2001-2009. **Stroke 43:**3200–3206, 2012
10. Center for Health Policy, Center for Primary Care and Outcomes Research, Battelle Memorial Institute: **Quality Indicator Measure Development, Implementation, Maintenance, and Retirement.** Rockville, MD: Agency for Healthcare Research and Quality, 2011 (http://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/QI_Measure_Development_Implementation_Maintenance_Retirement_Full_5-3-11.pdf) [Accessed May 1, 2014]
11. Charlson ME, Pompei P, Ales KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. **J Chronic Dis 40:**373–383, 1987
12. Claassen J, Bateman BT, Willey JZ, Inati S, Hirsch LJ, Mayer SA, et al: Generalized convulsive status epilepticus after nontraumatic subarachnoid hemorrhage: the nationwide inpatient sample. **Neurosurgery 61:**60–65, 2007
13. Cowan JA Jr, Dimick JB, Wainess RM, Upchurch GR Jr, Thompson BG: Outcomes after cerebral aneurysm clip occlusion in the United States: the need for evidence-based hospital referral. **J Neurosurg 99:**947–952, 2003
14. Crowley RW, Yeoh HK, Stukenborg GJ, Ionescu AA, Kassell NF, Dumont AS: Influence of weekend versus weekday hospital admission on mortality following subarachnoid hemorrhage. Clinical article. **J Neurosurg 111:**60–66, 2009
15. Daley J, Jencks S, Draper D, Lenhart G, Thomas N, Walker J: Predicting hospital-associated mortality for Medicare patients. A method for patients with stroke, pneumonia, acute myocardial infarction, and congestive heart failure. **JAMA 260:**3617–3624, 1988
16. Elixhauser A, Steiner C, Harris DR, Coffey RM: Comorbidity measures for use with administrative data. **Med Care 36:**8–27, 1998
17. Fleming ST, Hicks LL, Bailey RC: Interpreting the Health Care Financing Administration's mortality statistics. **Med Care 33:**186–201, 1995
18. Fonarow GC, Pan W, Saver JL, Smith EE, Reeves MJ, Broderick JP, et al: Comparison of 30-day mortality models for profiling hospital performance in acute ischemic stroke with vs without adjustment for stroke severity. **JAMA 308:**257–264, 2012
19. Freeman JL, Fetter RB, Park H, Schneider KC, Lichtenstein JL, Hughes JS, et al: Diagnosis-related group refinement with diagnosis- and procedure-specific comorbidities and complications. **Med Care 33:**806–827, 1995
20. George MG, Tong X, Kuklina EV, Labarthe DR: Trends in stroke hospitalizations and associated risk factors among children and young adults, 1995-2008. **Ann Neurol 70:**713–721, 2011
21. Goldman LE, Chu PW, Osmond D, Bindman A: The accuracy of present-on-admission reporting in administrative data. **Health Serv Res 46:**1946–1962, 2011
22. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP: Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. **JAMA 264:**2768–2774, 1990
23. Hoh BL, Chi YY, Lawson MF, Mocco J, Barker FG II: Length of stay and total hospital charges of clipping versus coiling for ruptured and unruptured adult cerebral aneurysms in the Nationwide Inpatient Sample database 2002 to 2006. **Stroke 41:**337–342, 2010
24. Hoh BL, Kleinhenz DT, Chi YY, Mocco J, Barker FG II: Incidence of ventricular shunt placement for hydrocephalus with clipping versus coiling for ruptured and unruptured cerebral aneurysms in the Nationwide Inpatient Sample database: 2002 to 2007. **World Neurosurg 76:**548–554, 2011
25. Hoh BL, Nathoo S, Chi YY, Mocco J, Barker FG II: Incidence

of seizures or epilepsy after clipping or coiling of ruptured and unruptured cerebral aneurysms in the nationwide inpatient sample database: 2002-2007. **Neurosurgery 69:**644–650, 2011
26. Hunt WE, Hess RM: Surgical risk as related to time of intervention in the repair of intracranial aneurysms. **J Neurosurg 28:**14–20, 1968
27. Jencks SF, Daley J, Draper D, Thomas N, Lenhart G, Walker J: Interpreting hospital mortality data. The role of clinical risk adjustment. **JAMA 260:**3611–3616, 1988
28. Kasner SE: Clinical interpretation and use of stroke scales. **Lancet Neurol 5:**603–612, 2006
29. Khatri R, Tariq N, Vazquez G, Suri MF, Ezzeddine MA, Qureshi AI: Outcomes after nontraumatic subarachnoid hemorrhage at hospitals offering angioplasty for cerebral vasospasm: a national level analysis in the United States. **Neurocrit Care 15:**34–41, 2011
30. Kim YW, Neal D, Hoh BL: Cerebral aneurysms in pregnancy and delivery: pregnancy and delivery do not increase the risk of aneurysm rupture. **Neurosurgery 72:**143–150, 2013
31. Kuklina EV, Tong X, Bansil P, George MG, Callaghan WM: Trends in pregnancy hospitalizations that included a stroke in the United States from 1994 to 2007: reasons for concern? **Stroke 42:**2564–2570, 2011
32. Kurth T, Elkind MS: Comparing hospitals on stroke care: the need to account for stroke severity. **JAMA 308:**292–294, 2012
33. Landis JR, Koch GG: The measurement of observer agreement for categorical data. **Biometrics 33:**159–174, 1977
34. Leake CB, Brinjikji W, Kallmes DF, Cloft HJ: Increasing treatment of ruptured cerebral aneurysms at high-volume centers in the United States. Clinical article. **J Neurosurg 115:** 1179–1183, 2011
35. Lin N, Cahill KS, Frerichs KU, Friedlander RM, Claus EB: Treatment of ruptured and unruptured cerebral aneurysms in the USA: a paradigm shift. **J Neurointerv Surg 4:**182–189, 2012
36. Mayes R: The origins, development, and passage of Medicare's revolutionary prospective payment system. **J Hist Med Allied Sci 62:**21–55, 2007
37. McDonald RJ, McDonald JS, Bida JP, Kallmes DF, Cloft HJ: Subarachnoid hemorrhage incidence in the United States does not vary with season or temperature. **AJNR Am J Neuroradiol 33:**1663–1668, 2012
38. McMahon LF Jr: Diagnosis-related groups: past and future. **Infect Control Hosp Epidemiol 9:**471–474, 1988
39. McMahon LF Jr, Newbold R: Variation in resource use within diagnosis-related groups. The effect of severity of illness and physician practice. **Med Care 24:**388–397, 1986
40. Mehta RH, Liang L, Karve AM, Hernandez AF, Rumsfeld JS, Fonarow GC, et al: Association of patient case-mix adjustment, hospital process performance rankings, and eligibility for financial incentives. **JAMA 300:**1897–1903, 2008
41. Molyneux A, Kerr R, Stratton I, Sandercock P, Clarke M, Shrimpton J, et al: International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised trial. **Lancet 360:**1267–1274, 2002
42. Nguyen-Huynh MN, Johnston SC: Regional variation in hospitalization for stroke among Asians/Pacific Islanders in the United States: a nationwide retrospective cohort study. **BMC Neurol 5:**21, 2005
43. Nuño M, Patil CG, Lyden P, Drazin D: The effect of transfer and hospital volume in subarachnoid hemorrhage patients. **Neurocrit Care 17:**312–323, 2012
44. Ovbiagele B: Nationwide trends in in-hospital mortality among patients with stroke. **Stroke 41:**1748–1754, 2010
45. Quinn TJ, Dawson J, Walters MR, Lees KR: Reliability of the modified Rankin Scale: a systematic review. **Stroke 40:**3393–3395, 2009
46. Qureshi AI, Vazquez G, Tariq N, Suri MF, Lakshminarayan K, Lanzino G: Impact of International Subarachnoid Aneurysm Trial results on treatment of ruptured intracranial aneurysms in the United States. Clinical article. **J Neurosurg 114:**834–841, 2011
47. Rosen DS, Macdonald RL: Subarachnoid hemorrhage grading scales: a systematic review. **Neurocrit Care 2:**110–118, 2005
48. Rosenthal GE, Baker DW, Norris DG, Way LE, Harper DL, Snow RJ: Relationships between in-hospital and 30-day standardized hospital mortality: implications for profiling hospitals. **Health Serv Res 34:**1449–1468, 2000
49. Shahian DM, Edwards FH: The Society of Thoracic Surgeons 2008 cardiac surgery risk models: introduction. **Ann Thorac Surg 88 (1 Suppl):**S1, 2009
50. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al: The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. **Ann Thorac Surg 88 (1 Suppl):**S2–S22, 2009
51. Shea AM, Reed SD, Curtis LH, Alexander MJ, Villani JJ, Schulman KA: Characteristics of nontraumatic subarachnoid hemorrhage in the United States in 2003. **Neurosurgery 61:**1131–1138, 2007
52. Shukla P, Fisher R, Fisher R: Testing of 3M's APR-DRG risk adjustment for hospital mortality outcomes. **Abstr Acad Health Serv Res Health Policy Meet 19:**11, 2002 (Abstract)
53. Siddiq F, Chaudhry SA, Tummala RP, Suri MF, Qureshi AI: Factors and outcomes associated with early and delayed aneurysm treatment in subarachnoid hemorrhage patients in the United States. **Neurosurgery 71:**670–678, 2012
54. Smits HL, Fetter RB, McMahon LF Jr: Variation in resource use within diagnosis-related groups: the severity issue. **Health Care Financ Rev Suppl:**71–78, 1984
55. Teasdale GM, Drake CG, Hunt W, Kassell N, Sano K, Pertuiset B, et al: A universal subarachnoid hemorrhage scale: report of a committee of the World Federation of Neurosurgical Societies. **J Neurol Neurosurg Psychiatry 51:**1457, 1988
56. US Health Care Financing Administration: **Medicare Hospital Mortality Information, 1987, 1988, 1989.** Washington, DC: Department of Health and Human Services, 1989
57. van Walraven C, Austin PC, Manuel D, Knoll G, Jennings A, Forster AJ: The usefulness of administrative databases for identifying disease cohorts is increased with a multivariate model. **J Clin Epidemiol 63:**1332–1341, 2010
58. World Federation of Neurological Surgeons Committee: Report of World Federation of Neurological Surgeons Committee on a Universal Subarachnoid Hemorrhage Grading Scale. **J Neurosurg 68:**985–986, 1988 (Letter)

Manuscript submitted May 7, 2013.
Accepted April 29, 2014.
Portions of this work were presented as an abstract at the American Heart Association/American Stroke Association February 2013 International Stroke Conference and Nursing Symposium in Honolulu, HI.
Please include this information when citing this paper: published online June 20, 2014; DOI: 10.3171/2014.4.JNS131100.
Address correspondence to: Chad W. Washington, M.D., M.S., M.P.H.S., Washington University School of Medicine, Department of Neurological Surgery, 660 S. Euclid Ave., Campus Box 8057, St. Louis, MO 63110. email: washingtonc@wudosis.wustl.edu.