Risk Adjustment of Ischemic Stroke Outcomes for Comparing Hospital Performance

A Statement for Healthcare Professionals From the American Heart Association/American Stroke Association

Irene L. Katzan, MD, MS, FAHA, Chair; John Spertus, MD, FAHA, Co-Chair;
Janet Prvu Bettger, ScD, FAHA*; Dawn M. Bravata, MD*;
Mathew J. Reeves, PhD, DVM, FAHA*; Eric E. Smith, MD, MPH, FAHA*;
Cheryl Bushnell, MD, MHS, FAHA; Randall T. Higashida, MD, FAHA;
Judith A. Hinchey, MD, FAHA; Robert G. Holloway, MD, MPH; George Howard, DrPH;
Rosemarie B. King, PhD, RN, FAHA; Harlan M. Krumholz, MD, FAHA;
Barbara J. Lutz, PhD, RN, CRRN, FAHA; Robert W. Yeh, MD, MSc, FAHA; on behalf of the
American Heart Association Stroke Council, Council on Quality of Care and Outcomes Research, Council on Cardiovascular and Stroke Nursing, Council on Cardiovascular Radiology and Intervention, Council on Cardiovascular Surgery and Anesthesia, and Council on Clinical Cardiology

Background and Purpose—Stroke is the fourth-leading cause of death and a leading cause of long-term major disability in the United States. Measuring outcomes after stroke has important policy implications. The primary goals of this consensus statement are to (1) review statistical considerations when evaluating models that define hospital performance in providing stroke care; (2) discuss the benefits, limitations, and potential unintended consequences of using various outcome measures when evaluating the quality of ischemic stroke care at the hospital level; (3) summarize the evidence on the role of specific clinical and administrative variables, including patient preferences, in risk-adjusted models of ischemic stroke outcomes; (4) provide recommendations on the minimum list of variables that should be included in risk adjustment of ischemic stroke outcomes for comparisons of quality at the hospital level; and (5) provide recommendations for further research.

Methods and Results—This statement gives an overview of statistical considerations for the evaluation of hospital-level outcomes after stroke and provides a systematic review of the literature for the following outcome measures for ischemic stroke at 30 days: functional outcomes, mortality, and readmissions. Data on outcomes after stroke have primarily involved studies conducted at an individual patient level rather than a hospital level. On the basis of the available information, the following factors should be included in all hospital-level risk-adjustment models: age, sex, stroke severity, comorbid conditions, and vascular risk factors. Because stroke severity is the most important prognostic factor for individual patients and appears to be a significant predictor of hospital-level performance for 30-day mortality, inclusion of a stroke severity measure in risk-adjustment models for 30-day outcome measures is recommended. Risk-adjustment models that do not include stroke severity or other recommended variables must provide comparable classification of hospital performance.
Evaluating patient outcomes is a central part of optimizing the quality of patient care. In an effort to improve healthcare quality, regulatory bodies have begun to measure outcomes after hospitalization for several diseases, including pneumonia, myocardial infarction, and congestive heart failure. Similarly, stroke is an emerging, high-priority target for such efforts. Stroke is the fourth-leading cause of death and a cause of long-term major disability in the United States. It is also one of the top 10 costliest conditions for the Centers for Medicare & Medicaid Services (CMS). As with most disease states, accurately measuring and comparing hospitals' outcomes is complex. Some outcome determinants are outside a provider's control, including patient factors such as age, the severity of illness, comorbid conditions, and preferences for particular treatments. Domains that are modifiable and potentially improvable include the content and organization of the stroke care delivery system. These include factors such as the relationship between providers within a particular community (eg, emergency medical services, hospitals, nursing support, rehabilitation services, and home environments), hospice availability, and the types and skills of providers. Hospitals also have control over the use of acute (eg, cerebral reperfusion) and chronic (eg, rehabilitation services) treatments. The challenge in using outcomes as a measure of healthcare quality is to account for the unmodifiable factors so that variations in outcomes between providers are attributable to the processes of stroke care and more accurately reflect the quality of care provided.

A clear link between quality of hospital care and outcomes after stroke has not been well demonstrated in the literature. The concept that differences in case-mix-adjusted outcomes reflect quality has also been challenged. However, given the intense interest in systematically measuring stroke outcomes and in comparing outcomes across different facilities so that the quality of stroke care can be evaluated and optimized, there is a need to define the relevant stroke outcomes to measure, as well as the factors that should be accounted for to make comparisons across different facilities valid. The present scientific statement strives to illuminate the benefits and limitations of various stroke outcome measures and to provide guidance on the factors that should be used in risk adjustment so that these outcome measures will be most useful to clinical scientists, clinicians, policy makers, and quality experts when evaluating the quality of stroke care.

Specific goals of the statement are as follows:

- Discuss the benefits, limitations, and potential unintended consequences of using various outcome measures, including 30-day mortality, 30-day all-cause readmission, and 30-day functional status, when evaluating the quality of stroke care at the hospital level.
- Summarize the evidence on the role of specific clinical and administrative variables, including patient preferences, in risk-adjusted models of stroke outcomes.
- Provide recommendations on the minimum list of variables that should be included in risk adjustment for stroke outcomes when quality is compared at the hospital level.
- Provide recommendations for further research.

The present statement provides a systematic review of the literature that focuses on outcome measures for ischemic stroke. It does not include measures for hemorrhagic stroke or transient ischemic attack because of differences in management and outcomes for these conditions. The writing group chose the 30-day time period as a reasonable compromise for a standardized framework for measuring outcomes from the hospital's perspective, while recognizing that some outcomes, such as functional status, continue to improve over time.

Methodology and Evidence Overview

The Stroke Council of the American Heart Association (AHA)/American Stroke Association commissioned the assembled authors, who represent the fields of vascular neurology, biostatistics, epidemiology, health service research, nursing, rehabilitative medicine, and neuroradiology. The writing committee was organized into sections, which were led by section leaders. Literature review was performed and initial drafts were written within each section. Data compilation and recommendation was discussed and revised by the full committee over a series of teleconferences that took place over 9 months.

Search criteria included articles published in English in 1980 through 2011 that were indexed in MEDLINE (PubMed). Because of potential differences in healthcare practices and outcomes across countries, this review included only studies from high-income countries, using a list provided by the World Bank. Studies that included placebo groups from clinical trials were reviewed because they may provide systematically collected and relevant data on outcomes of interest. We were cognizant of the potential limitation of reduced generalizability of data from such highly selected patients. Search and article reviews were limited to studies of ischemic stroke patients. Although the focus of the present statement is on 30-day outcomes, we also included studies that evaluated mortality at the time of discharge to avoid missing variables.
in these studies found to be strongly associated with mortality. In addition, we expanded our review to include analyses of functional outcomes at 90 days, because studies have shown that function and disability at 30 days are highly predictive of function at 90 days.7,8

In the review of functional outcome measures, studies that used a composite end point for “poor” outcomes that combined death with severe disability or dependence were excluded. Combining poor function and death prevents the ability to discern the variables that predict poor function separately from those that predict mortality and inappropriately equates limitations in activities of daily living with death.9 From a quality perspective, it is likely that acute interventions that have no effect on mortality may affect function, and these relationships may not be detected with the use of a composite end point of disability and death.10

Results

Statistical Considerations
Differences between hospitals in the mortality, readmission, and functional status of stroke patients can be attributable to “hospital effects” that occur because of differences in the quality of hospital care, or to “case-mix effects” that occur because of differences in the characteristics of patients treated, or to random chance. When seeking to quantify hospital quality, the most common application of risk-adjustment methods is to use multivariable models to mitigate the contribution of measured patient characteristics present at the time of admission, such as differences in age, stroke severity, comorbidities, and risk factors. This assumes, although it is frequently unstated, that the remaining variation in outcomes at the hospital level represents the direct effect of differences in the quality of care delivered by the hospitals, after random error is taken into account.11 However, studies have not always shown associations between hospital risk-adjusted outcomes and poorer-quality care.5,6 Possible explanations for the inconsistent relationships mostly relate to fundamental limitations of observational study designs,12 but shortcomings in the underlying quality metrics to reflect changes in disease outcomes should also be considered. We focus, in this section, on the use of risk-adjustment methods to mitigate patient case-mix effects so as to be able to compare outcomes across different hospitals for the purpose of profiling, rather than on generating individual patient-level prognostic models.13 The latter approach to risk prediction at the individual level14 is important to both the patient and physician but is best viewed as a different application of these methods.

Methodological Standards for Risk-Adjustment Models Used for Hospital Profiling
Krumholz and colleagues11 have previously summarized standards for the development of risk-adjustment models when used for public reporting of healthcare providers’ outcomes (Table 1). Issues that are especially relevant to stroke include the following:

1. Sufficiently high-quality and timely data: Data can be identified from administrative databases, medical records, or specific efforts to collect quality of care data prospectively, such as Get With the Guidelines (GWTG)–Stroke.15 Administrative data are collected primarily for billing purposes and may not capture important clinical data known to be associated with outcomes.16 A significant challenge for risk adjustment of stroke outcomes with either administrative or clinical data is that measures of stroke severity, such as the National Institutes of Health Stroke Scale (NIHSS), are frequently missing.17 Moreover, these missing data are typically not missing at random, which has the potential to introduce important selection biases.

2. Consistency in coding across sites: A further complication of using administrative data is that variation across hospitals in the coding of comorbidities and risk factors can have an important impact on risk-adjustment models and the accuracy of hospital profiling. Consistency of coding can refer to the variation between hospitals in how they define the presence or absence of a specific comorbidity or risk factor. An example of this phenomenon is the coding of shock in heart failure or myocardial infarction patients. Shock is an important predictor of outcomes in these patients, but there are differences in the coding of shock across centers that may represent differences in the definition of shock rather than variation in its underlying prevalence.

Consistency of coding can also refer to variation in the number of comorbidities and risk factors recorded by a hospital for a given patient. Some hospitals routinely code for a larger number of comorbidities and risk factors than other hospitals, and this “deeper” level of coding can have important implications for the accuracy of the risk-adjustment process. Variation in the depth of coding has been shown to produce paradoxical increases in the very differences for which the risk-adjustment models are attempting to adjust.18,19

3. Designation of an appropriate reference time before which covariates are derived and after which outcomes are measured: The goal of risk adjustment is to adjust for the condition of the patient at the time patient care began (referred to as the reference time). This requires identifying and distinguishing between historical comorbid conditions that may have resolved at the time of admission (which do not need to be adjusted for in
the analysis), comorbid conditions present at the time of admission (which do need to be adjusted for), and complications or other events that developed during or after the admission (for which no adjustment should be made). The inclusion of factors that occurred after the reference time may “adjust away” the differences in quality of care, thereby defeating the central purpose of the model itself. A limiting factor in risk-adjustment models that use administrative data or clinical registries is that they typically do not include information regarding the timing of complications or comorbid conditions in relation to the stroke event.

4. Use of an appropriate outcome and a standardized period of outcome: Any outcome chosen as a measure of quality of care should meet 2 fundamental standards. First, the outcome should reflect an entity of interest to patients. Second, evidence should exist to support the relationship between the outcome and quality of care. The time period over which the outcome is assessed also needs to be standardized. In clinical stroke trials, 90 days tends to be the standard time period used for the evaluation of functional outcomes, whereas the AHA/American College of Cardiology task force on performance measures has recommended that stroke outcomes be measured at 30 days, which is also the time frame that CMS has adopted for a majority of its publicly reported outcome measures. We support the use of the shorter 30-day time period because it reduces the potential for loss to follow-up and mitigates the effects of differences in care that occur after hospitalization. Outcomes measured at hospital discharge may have some utility because they are easier to collect, but they are limited by differences in length of stay and the effects of patient transfers between hospitals. In addition, outcomes measured at discharge are insufficient to assess improvement in functional recovery, which typically occurs over several weeks. It has been recommended that discharge measures should only be used when validated against a standardized period of assessment. We recognize that different opinions on timing of outcomes measurement exist, and we suggest more research be conducted to understand the impact of using in-hospital measures as proxies of the more desirable, but often impractical, longer-term outcomes.

Model Development and Application of Risk-Adjustment Models for Hospital Profiling

Selection of variables for inclusion in risk-adjustment models should follow the previous published guidelines by Harrel et al and Iezzoni. The goal is to develop a model that captures the patient information that is significantly related to the outcome of interest. For risk-adjustment modeling to accurately identify hospital-level effects, the model must include strong and valid patient-level predictors that accurately stratify risk at the hospital level. It is therefore important to recognize that if a strong patient-level predictor variable is distributed similarly across different hospital populations, then controlling for it in a model will have little effect on the risk-adjusted hospital-level outcomes; only factors that are both related to the outcome and are differentially distributed between hospitals can act as confounders in the assessment of hospital differences. For example, a study of US Department of Veterans Affairs (VA) hospitals found that although the NIHSS was a strong predictor of mortality among individual ischemic stroke patients, it had little impact on the calculated risk-adjusted standardized mortality rates, probably because variation in both NIHSS and 30-day mortality was limited across the VA hospitals in the study.

The traditional statistical approach used for hospital profiling involves the application of risk-adjustment models to generate hospital-specific risk-standardized ratios that are used to rank order providers and to identify “outlier” hospitals. These ratios are generated by comparing the observed event rate at a specific hospital with the expected event rate, which is calculated by applying the regression coefficients from the risk-adjusted model to the hospital’s specific patient population. An observed/expected ratio >1 indicates higher than expected events for a particular hospital, whereas a ratio <1 indicates fewer than expected events. Often the observed/expected ratio is multiplied by the average event rate in the underlying population, which generates a risk-standardized outcome rate for each hospital. This approach, referred to as indirect standardization, has been heavily criticized for failing to account for variation in case numbers across hospitals or for intrahospital clustering effects.

Newer methods based on hierarchical or multilevel models that use random effect terms to describe hospital-specific effects address these deficiencies and are now the standard approach used for hospital profiling by CMS. The first level of a hierarchical model includes the patient characteristics, whereas the second level includes hospital-specific random intercepts that allow for different baseline event rates between the hospitals. The hospital-specific intercepts also account for the clustering of patients within each hospital. Somewhat confusingly, random effects models are described as generating predicted versus expected ratios, rather than the observed versus expected ratios described previously. Using mortality as an example, the expected number of deaths at a hospital from a random effects model is obtained by applying the model regression coefficients (generated from all the patients in the sample) to the patient population observed in the particular hospital; adding the average of the hospital-specific intercepts; transforming; and then summing over all patients in the hospital to get the estimated number of deaths. The predicted number of deaths is calculated similarly except that the hospital-specific random intercept (which represents the baseline mortality rate of the particular hospital) is substituted for the average hospital-specific intercept. The predicted over expected ratio is often multiplied by the unadjusted mortality rate for the total population to yield a risk-standardized mortality rate. Other potential standardization approaches include generalized estimating equations, which can account for clustering without invoking random effects.

Assessment of Model Fit (Calibration) and Discrimination

There are 2 existing, albeit imperfect, methods for describing the accuracy of a risk-adjustment model at the patient-level: calibration and discrimination.

Calibration describes the degree to which a model’s predicted probabilities match the observed data. In a perfectly
calibrated model, a predicted risk estimate of x% means that exactly x% of patients are observed to have the outcome. Discrimination is measured by the C statistic. The closer the C statistic is to 1.0, the better the model can predict outcomes at an individual patient level. An important limitation of the C statistic is that it only measures the predictive power of a model at the patient level and does not directly express the ability of the model to accurately profile hospitals with respect to the hospital-specific risk-standardized ratios. Both calibration and discrimination statistics are important measures to be evaluated when judging the performance of risk-adjustment models.

Further Methodological Issues in the Development of Risk-Adjusted Models for Stroke

Limited Sample Size

Even with complete adjustment for all important prognostic factors, differences between hospitals may be caused by chance alone. Small hospitals, by definition, provide fewer data, and so the precision of model estimates is lower; estimates are made with more uncertainty, which is reflected by wider confidence intervals. The frequency of the outcome being modeled also impacts relative precision. Outcomes such as 30-day case fatality and 30-day readmission after stroke are not that common; unadjusted 30-day event rates for both mortality and readmission are typically around 12% to 15% depending on the population examined. Small case volumes in combination with low frequency of outcome events have a significant negative impact on the ability to identify hospitals as outliers (as is discussed in Hospital Profiling: Identification of Outlier Hospitals). Statistical methods such as hierarchical or multilevel regression models and empirical Bayes methods are increasingly being used to address the problem of risk adjustment across small hospitals or for health events that are infrequent. Consideration of small case volume is important in risk-adjustment models for stroke, because some studies have shown a relationship between stroke volumes and outcomes.

Quantification of Outcomes as a Graded Response

Functional outcome measures, such as the modified Rankin Scale (mRS), are recorded on an ordinal scale. Several statistical approaches are available for the analysis of such data, such as a binary logistic model that uses a fixed dichotomous cut point (eg, <2 versus ≥2), a sliding dichotomy (or responder analysis), or an ordinal analysis. Each method has its own strengths and limitations, and no one method is always the preferred choice. The reader is referred to 2 recent summaries for further information on this issue.

Hospital Profiling: Identification of Outlier Hospitals

Hospital-specific risk standardized ratios are commonly used to rank order hospitals from highest to lowest performers. Many quality improvement organizations and payers such as CMS also use hospital-specific risk-standardized estimates to identify outlier hospitals. Outlier hospitals may be identified simply according to relative rank (eg, top 10%, bottom 10%) or by identifying hospitals that are statistically significantly different from a given benchmark or standard (commonly the overall average mortality or readmission rate in the population being evaluated). It is important to appreciate that unlike an assessment of the performance of risk-adjustment models at the patient level, in which the final outcome of each patient (eg, alive or dead, readmitted or not) is known, the true rank order of hospitals that describes their relative performance is not known. Because the designation of outlier status can have important consequences for individual hospitals, the accuracy of such designations has been of keen interest to researchers in the field. Findings from this research are quite sobering. In 2 seminal simulation studies conducted in the 1990s, it was found that even in the face of perfect case-mix adjustment and very large variations in quality between hospitals, mortality rates did a very poor job of identifying low-quality hospitals, identifying only a small minority of poorly performing hospitals and generating far more false-positive than true-positive results. These studies also illustrated that the ability to identify outlier hospitals was strongly influenced by hospital case volumes and the underlying mortality rates. Other reports have demonstrated that the accuracy of hospital profiling can be influenced by the statistical methods used (for example, fixed versus random effects models, Bayesian versus frequentist methods, the choice of modeling strategy used to identify outlier or extreme providers, and the type and quality of data available).

Thirty-Day Mortality:

Strengths and Weaknesses of 30-Day Mortality as a Measure of Hospital-Level Quality of Care in Ischemic Stroke

There are several reasons to consider the use of mortality as a measure of quality in ischemic stroke care. First, death is obviously an important outcome. Second, death is easily measured and reliably assessed through administrative data, at little expense. Third, death after stroke is not uncommon, occurring in 5% to 7% of cases in the acute care hospital, in 13% to 15% at 30 days, and in 25% to 30% at 1 year. Fourth, there is variation in mortality rates between hospitals, which could be related to the quality of care. The variation in hospital ischemic stroke mortality occurs despite adjustment for age, sex, and comorbid conditions. For example, variations in hospital mortality have been associated with hospital size, the timing of admission, and aggressiveness in end-of-life care, all of which may be associated with the quality of care provided.
In a study of Medicare ischemic stroke patients admitted to 625 hospitals participating in a quality improvement program, hospital 30-day risk-adjusted standardized mortality varied by >75% on a relative basis between the bottom tenth percentile (9.8%) and the top tenth percentile (17.8%). Even after adjustment for the NIHSS, similar variability was observed. Fifth, better organized stroke care is associated with lower mortality. For example, organized inpatient ischemic stroke care lowered the odds of mortality by 14% compared with routine care according to a systematic review of controlled trials. An observational study showed an absolute mortality reduction of 2.5% in state-certified stroke centers compared with noncertified centers. Beyond such structural measures, processes of care are also associated with improved mortality. Aspirin use in the first 48 hours has an impact on longer-term (eg, 6 months) mortality, albeit small, and there is some evidence that the use of dysphagia screens, prophylaxis for deep venous thrombosis, and treatment of hypoxia decreases the risk of death or discharge to hospice.

There are also important limitations or challenges to the use of mortality as a measure of the quality of care in ischemic stroke. First, the outcome of patients with stroke, including mortality, is associated with the initial severity of the stroke, age, and comorbidities, none of which are modifiable, although they can be adjusted for in risk models. Second, the degree to which ischemic stroke mortality is modifiable by changes in processes is unclear. The only medical or surgical interventions proven by randomized controlled trials to lower the risk of death in ischemic stroke are early provision of aspirin, which has a minor effect at 6 months, and hemicraniectomy for malignant cerebral edema, which has a large effect but is applicable to only a small fraction of patients. The biggest advance in acute stroke treatment in the past 20 years is the use of thrombolytic therapy, yet thrombolysis with tissue-type plasminogen activator does not reduce mortality (although it does prevent disability). Therefore, although organized inpatient care appears to lower mortality, the mechanism by which it does so is unclear. Third, 30-day mortality reflects not only the care provided by the acute admitting hospital but also the care provided immediately after acute discharge, including rehabilitative care, although hospitals’ use of such services and their transitions from inpatient to outpatient care can be an important opportunity for quality improvement. Finally, short-term stroke mortality may be more reflective of patient/family preferences than the provision of evidence-based care. An earlier death may be preferable to some patients and their families over a delayed death or prolonged state of severe disability, depending on the patient’s life stage, functional and cognitive status, treatment preferences, and desired place of death. As a result, life-sustaining treatments are sometimes withheld in accordance with patients’ preferences, which leads to early death. Therefore, hospitals that practice optimal shared decision making could have higher mortality rates, caused by higher rates of appropriate use of palliative care, than centers that are less sensitive to patients’ preferences. Because the actual patient and family preferences and values (and the process of eliciting them) may be difficult to capture, this can limit the interpretability of mortality comparisons, and it underscores the importance of a broader range of processes and outcome measures, including functional status, for use in quantifying the quality of stroke care.

**Review of Evidence on Models That Predict Mortality After Ischemic Stroke**

The systematic review identified 1168 articles related to prediction of mortality after ischemic stroke, of which 10 were deemed to have the most relevant information on individual predictors of risk of death (Table 2). Key features of the most relevant models are presented in Table 2. The reported discrimination, as measured by the C statistic, varied from 0.72 to 0.86, with most models having C statistics ≥0.80, which suggests that ischemic stroke mortality can be predicted at the patient level with a degree of accuracy that is often considered clinically meaningful.

Most of the published models have design limitations that make them inadequate for the purpose of hospital-level ischemic stroke mortality risk adjustment, however. Most were designed to predict an individual’s risk of death after ischemic stroke and did not report the impact of risk adjustment on hospital-level mortality rates. Some models predicted hospital discharge mortality rather than mortality at a fixed time point, such as 30 days. Some used postadmission information that could reflect the quality of care provided or included patients in the subacute phase of stroke. Some models were derived or validated in selected populations, such as patients participating in clinical trials, or have not been validated in new independent populations, which potentially limits their generalizability. Only some models excluded transfer-in patients, although the inclusion of transferred patients could result in double counting of the same stroke.

Despite the heterogeneity in patient populations and methods of the previously published models (Table 2), some consistent strong predictors of death after ischemic stroke are evident at the patient level. Measures of stroke severity, either by a standardized stroke scale score such as the NIHSS or by clinical examination features, were consistently retained in models for which such information was available. Likewise, age was consistently included. Medical comorbidities were included in many models, although the exact comorbidities in each model varied. In studies that compared the predictive value of various types of predictors, stroke severity and age explained the greatest variation in predicting mortality at the patient level and substantially increased the accuracy of the models. In a study of in-hospital mortality, the effect of the NIHSS on reclassifying risk was assessed by use of the integrated discrimination index, which quantifies the difference between 2 models with and without the candidate predictor of interest, in the predicted risk of death in case subjects (ie, those who died) versus control subjects (those who survived). A higher integrated discrimination index indicates improved discrimination of one model over another. The integrated discrimination index for the addition of the NIHSS to the model for prediction of in-hospital mortality was 9.4%, which indicates a substantial improvement in classification of risk. In a study of Medicare beneficiaries, there was a graded, nearly linear relationship between higher NIHSS and higher 30-day mortality. The discrimination of risk of 30-day mortality by NIHSS alone (C statistic=0.82) was better than that of a model that included...
The role of laboratory values or physiological findings such as blood pressure in predicting ischemic stroke death is less certain, with fewer data than for stroke severity, age, and medical comorbidities. A recent systematic review found that most existing studies of the relationship between

### Table 2. Ischemic Stroke Mortality Prediction Models

<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Setting (Derivation Sample)</th>
<th>Sample Size (Derivation), n</th>
<th>Mortality Outcome</th>
<th>Significant Prognostic Factors</th>
<th>Validation</th>
<th>C Statistic (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counsell et al</td>
<td>2002</td>
<td>UK population based, including outpatients</td>
<td>530</td>
<td>30 d</td>
<td>Age, living alone, independent before stroke, Glasgow Coma Scale verbal score, ability to lift arms and ability to walk</td>
<td>Yes (in 3 stroke registries in Scotland, Italy, and Australia with 1868 patients)</td>
<td>0.85</td>
</tr>
<tr>
<td>Hinchey et al</td>
<td>1998</td>
<td>31 Cleveland (OH) hospitals</td>
<td>223</td>
<td>Discharge</td>
<td>Altered mental status, limb paralysis, neurological examination findings, laboratory findings, CT or MR findings, vital signs, time of symptom onset, intubation, ECG findings, DNR order within first 3 d</td>
<td>No</td>
<td>Not reported</td>
</tr>
<tr>
<td>Iezzoni et al</td>
<td>1996</td>
<td>94 US hospitals</td>
<td>9407</td>
<td>Discharge</td>
<td>Compared different sets of predictors: (1) Medical record–collected age, demographics, diagnoses, procedures, examination findings, vital signs, and laboratory values; (2) physiology score based on vital signs; and (3) 3 vendor-specific disease severity measures based on discharge abstracts</td>
<td>Yes (internal only)</td>
<td>Medical record: 0.86; physiology score: 0.84; discharge summaries: 0.72–0.75</td>
</tr>
<tr>
<td>Lewis et al</td>
<td>2008</td>
<td>Randomized controlled trial</td>
<td>537</td>
<td>30 d</td>
<td>External validation of study by Counsell et al</td>
<td>Yes</td>
<td>0.73</td>
</tr>
<tr>
<td>Saposnik et al</td>
<td>2011</td>
<td>12 Canadian hospitals</td>
<td>8223</td>
<td>30 d</td>
<td>Age, Canadian Neurological Scale, sex, stroke subtype according to OCSP classification, atrial fibrillation, CHF, previous myocardial infarction, smoker, cancer, renal dialysis, preadmission dependency, serum glucose</td>
<td>Yes (population-based sample of 3270 hospitalized ischemic stroke patients from Ontario, Canada)</td>
<td>0.79</td>
</tr>
<tr>
<td>Smith et al</td>
<td>2010</td>
<td>1036 US hospitals</td>
<td>164993</td>
<td>Discharge</td>
<td>Age, NIHSS, mode of arrival to hospital, sex, atrial fibrillation, previous stroke or TIA, coronary artery disease, history of carotid stenosis, hypertension, diabetes mellitus, smoking, history of dyslipidemia, arrival during weekday working hours</td>
<td>Yes (109 995 separate patients at the same 1036 hospitals)</td>
<td>Including NIHSS (33.7% of patients): 0.85; not including NIHSS: 0.72</td>
</tr>
<tr>
<td>Tabak et al</td>
<td>2007</td>
<td>71 US acute care hospitals</td>
<td>89129</td>
<td>30-d hospital risk-adjusted mortality</td>
<td>Age, ICD-9 codes, vital signs, laboratory values, altered mental status</td>
<td>Yes (195 other US hospitals), but results were not reported separately except that they were “similar”</td>
<td>Age+laboratory findings: 0.75; age+laboratory findings+ICD-9 variables: 0.76; age+laboratory findings+ICD-9 variables+altered mental status: 0.84</td>
</tr>
<tr>
<td>Wang et al</td>
<td>2003</td>
<td>Single Australian hospital</td>
<td>253</td>
<td>1 y</td>
<td>Unconsciousness, dysphagia, urinary incontinence, both sides affected, hypertension, ischemic heart disease, peripheral vascular disease, diabetes mellitus</td>
<td>Yes (217 patients from same hospital)</td>
<td>Not reported</td>
</tr>
<tr>
<td>Weimar et al</td>
<td>2002</td>
<td>23 German stroke centers</td>
<td>1754</td>
<td>100 d</td>
<td>Age, NIHSS at admission, fever within 72 h</td>
<td>No</td>
<td>Not reported</td>
</tr>
<tr>
<td>Weimar et al</td>
<td>2004</td>
<td>7 German stroke centers</td>
<td>1079</td>
<td>100 d</td>
<td>Age, NIHSS</td>
<td>Yes (13 German hospitals separate from the derivation cohort)</td>
<td>0.83 (Derivation cohort)</td>
</tr>
</tbody>
</table>

CHF indicates congestive heart failure; CT, computed tomography; DNR, do not resuscitate; ICD-9, International Classification of Diseases, 9th Revision; MR, magnetic resonance; NIHSS, National Institutes of Health Stroke Scale; OCSP, Oxford Community Stroke Project; TIA, transient ischemic attack; and UK, United Kingdom.
blood markers and ischemic stroke were of relatively low quality, without a clear consensus on highly predictive variables.85

Critical issues in the consideration of risk models for mortality after acute ischemic stroke are whether or not clinical assessment of stroke severity is necessary and, if so, whether it is feasible. The importance of these issues has been heightened by CMS’s recent decision to publicly report stroke mortality using administrative data for risk adjustment, which does not include a measure of stroke severity. A mortality risk model was developed by the Yale-New Haven Health System/Center for Outcomes Research and Evaluation that included 42 variables from the Medicare claims files from the index hospitalization and from principal and secondary diagnosis codes from hospitalizations, institutional outpatient visits, and physician encounters in the 12 months before the index hospitalization. It also included a variable to account for transfers from an outside emergency department.86 In cases of hospital transfers after admission, mortality is attributed to the first hospital. The discrimination of the CMS model has been compared with a medical record–based model that included a stroke severity scale developed by the National Stroke Project that consisted of retrospective chart abstraction of the presence or absence of (1) vision changes, (2) speech deficit, (3) motor deficit, and (4) sensory deficit.87 The C statistic was 0.71 for the CMS model and 0.80 for the comparison medical record–based model. The CMS and medical record–based model predictions showed moderately good correlation (intraclass correlation coefficient 0.75). Analysis of differences in hospital rankings between the 2 models was not performed.

The National Quality Forum considered a performance measure for ischemic stroke mortality using the CMS/Yale risk-adjustment model; however, the performance measure was withdrawn before a decision on endorsement could be provided. A separate measure for stroke in-hospital mortality from the Agency for Healthcare Research and Quality has been approved by the National Quality Forum since 2008.88

Although adjustment for stroke severity is clearly the most important individual prognostic factor for stroke, a debated question is whether adjustment for stroke severity would change the order of ranking of hospital mortality from lowest to highest. If the distribution of severity scores at each hospital is similar, then hospital risk-adjusted mortality rates may not change much, although individual patient-level predictions would be improved. A study of VA hospitals found that adjustment for NIHSS caused little shift in hospital risk-adjusted mortality;6 however, there was little variability in hospital-specific risk-standardized mortality rates across VA hospitals, which may have limited the ability to detect differences with inclusion of the NIHSS in the models. Indeed, extrapolation of VA study results to the US population would suggest that the risk-standardized mortality rate for 30-day mortality is not able to differentiate hospital performance well; however, differences between the VA stroke population and the overall stroke population or Medicare-served population may limit the generalizability of the findings. The VA population was predominantly male, and approximately half had an NIHSS score ≤2, which reflects mild stroke. It also had a very low 30-day mortality rate (5.4%)76 compared with Medicare stroke patients (14.2%)77 or a population-based study in Canada (11.6%).78 which probably reflects preferential triage of acute stroke patients to non-VA hospitals. A recent study of hospitals participating in the AHA/American Stroke Association’s GWTG-Stroke quality initiative also evaluated the impact of NIHSS on hospital performance rankings.60 It included 782 hospitals and 127,950 Medicare beneficiaries with ischemic stroke who had a documented NIHSS. In this analysis, hospital transfers after admission were excluded, whereas the CMS analysis included hospital transfers but adjusted for them. There was a significant reclassification of hospitals when NIHSS was used in the model. Among 782 hospitals ranked into 3 groups (top 20%, middle 60%, or bottom 20%) of performance by the claims model without NIHSS scores, 26.3% (n=206) were ranked differently by the model with NIHSS scores. When classified on the basis of expected performance as defined by the CMS Hospital Compare program (based on having a 95% credible interval for the hospital-specific random effect that did not include the null value of zero, representing the overall hospital average), the NIHSS model reclassified 45 of 782 hospitals (5.7%), including 15 of 26 hospitals that were classified as “worse than expected” in the model without NIHSS. Although this report provides important information on the impact of including the NIHSS in risk-adjustment models for 30-day mortality, some unanswered questions remain. A limitation of the GWTG study is that 55% of patients were excluded because the NIHSS was not documented in the clinical record, and it is unknown whether inclusion of data from these patients would have changed the findings. Although the GWTG model-building strategy was designed to be similar to that used to derive models used by CMS for public reporting of ischemic stroke, the models do differ in some respects: There were more variables retained in the GWTG model than the CMS model; the GWTG model was derived and validated in non-transferred patients, whereas the CMS model included transfer patients and adjusted for transfer as another variable; and the GWTG study population was based on clinical diagnoses confirmed by chart review, whereas the CMS model was derived and validated on the basis of cases identified by administrative billing codes. Therefore, although the GWTG report60 provides useful information on the added value of the NIHSS for risk adjustment, it cannot be considered a direct validation of the model used by CMS for public reporting.

If stroke severity information were readily available at little expense, then clearly it would be desirable to adjust for it; however, standardized stroke scales are not included in administrative data and are not routinely recorded in clinical practice at all hospitals. Even among those participating in the national GWTG-Stroke quality improvement program, the NIHSS was recorded in only 40% of ischemic stroke patients.60 The feasibility of collecting the NIHSS or a simpler indicator of stroke severity at all hospitals in the United States has not been tested, and there has been no policy incentive to collect it to date. A performance measure to document admission stroke severity by use of a designated stroke scale would provide such motivation. In addition, we know little about the consistency of the quality of information from NIHSS scores across sites. In US practice, the NIHSS has become the preferred stroke scale and is recommended over other scales by
AHA/American Stroke Association guidelines. However, other validated stroke scales exist. A simpler stroke scale, such as the 3-item scale developed by Singer and colleagues, if validated, may be easier to implement nationally. The discriminative power for mortality risk prediction and feasibility in practice of these other scales compared with the NIHSS has not been tested comprehensively.

Given the current lack of widely available stroke severity information, an important question is whether a model without a severity measure can be used as a reasonable surrogate for a preferable model that includes a severity measure. Adequate surrogacy can be operationalized by determining whether the output of the model without a stroke severity measure provides comparable risk discrimination and classification of hospital performance as a model that includes severity. There was controversy within the writing committee, however, about whether this has been demonstrated adequately to date.

**Conclusions**

Conclusions on the minimum requirements for risk-adjustment data when mortality is used as an evaluation of stroke care quality at the hospital level are as follows:

1. Given our review of data on 30-day mortality models that primarily involved studies of predictors of 30-day mortality at the patient level, the following is suggested as a minimum list of variables that should be considered for inclusion: age, sex, stroke severity, comorbid conditions, and vascular risk factors.
2. Stroke severity is the most important prognostic factor for individual patients and appears to be a significant predictor in hospital-level performance. Inclusion in the prediction model is therefore recommended, particularly if it can be captured among all patients and reliably recorded by all hospitals.
3. Implementation without a measure of stroke severity will increase the occurrence of hospital misclassification.
4. Risk-adjustment models of stroke mortality that do not include stroke severity or other recommended variables must provide comparable classification of hospital performance as a model that includes stroke severity or other missing variables.
5. Stroke severity and other variables that are used in hospital-level risk-adjustment models of mortality after ischemic stroke should be standardized across sites so that their reliability and accuracy are equivalent.

The following represent needs for further research with mortality used as an indicator of hospital quality, in order of priority:

1. Research is urgently needed on the impact of hospital-level risk adjustment of stroke severity, comorbid conditions, and patient demographics on ischemic stroke mortality rates and their influence on hospital ranking by mortality, as well as the implications of misclassification for patients and hospitals, especially referral centers.
2. Research is needed to evaluate the impact of missing stroke severity variables on overall hospital performance.
3. More research is required to assess methods for more reproducibly and feasibly reporting stroke severity, or a validated close surrogate, in all ischemic stroke patients, suitable for incorporation into risk-adjustment models. For example, research is needed on alternative approaches to identifying stroke severity, such as using automated algorithms to search standardized clinical data stored in electronic health records.

4. More research is needed to identify hospital structures/organizations and processes of care that are associated with lower mortality. These processes could then be the targets for focused quality improvement interventions to reduce mortality.
5. Future work should be directed toward understanding how hospital and physician practices related to end-of-life treatment decisions impact 30-day stroke mortality. Ideally, methods for the capture of patient (not physician/hospital) preferences should be developed so that this information can be taken into consideration when hospital risk-adjusted mortality rates are calculated.
6. More research is needed on how differences in stroke systems of care, such as the capacity of providers and the beds available within a community (home, nursing, hospice), can impact 30-day stroke mortality.
7. More research is needed on the descriptive epidemiology of how stroke patients are actually dying in real world contemporary settings, with efforts to develop a typology aligned with measuring and improving the quality of care and identifying avoidable deaths.

Thirty-Day All-Cause Readmissions

Strengths and Weaknesses of Using 30-Day Readmission as an Outcome for Evaluation of Hospital Quality of Care

There are 4 key strengths to the use of 30-day all-cause readmission as a measure of hospital quality of care. First, it is patient-centered in that patients who experience this outcome incur the disruption, risk, and indirect (and sometimes direct) costs of the hospitalization and the clinical events that led to it. Second, as an outcome, readmission may serve as the end result of many distinct processes of care (some known and some unknown) that collectively combine to enhance or undermine a successful transition from hospital to posthospital care. Third, it is presumed, but unproven, that many readmissions could be prevented if care were improved. Hospital care involves the risks associated with multiple handoffs between providers and, commonly, discontinuity with outpatient providers. Research has shown that readmission rates for some diseases can be influenced by the quality of inpatient and outpatient care and that improvement in care coordination can reduce readmission rates. A number of studies, many of them in patients with heart failure, have demonstrated that improvements in care at the time of patient discharge can reduce 30-day readmission rates. Finally, readmissions are costly, and a reduction in these events would not only enhance the patient experience but could also reduce healthcare spending. Estimates suggest that potentially preventable readmissions may cost Medicare $12 billion annually. Interventions to reduce readmissions have been shown to decrease costs. For example, in a randomized study at an urban, academic, safety-net hospital, the intervention group that received discharge coordination, education, and
follow-up services had lower hospital utilization rates within 30 days of the index discharge than the control group that received usual discharge care, resulting in 33.9% lower costs for the intervention group.98

Balancing these 4 factors that support the use of readmission rates to evaluate hospital performance are 3 key limitations. First, not all readmissions are preventable, and so the goal will never be zero, but the premise is that the current rate includes many readmissions that could be prevented by improving care. Because it is difficult to determine the preventability of a specific readmission, it is typical to use a measure of all-cause readmission rather than a more specific metric of preventable readmission. Measurement of all-cause readmission should exclude planned or elective readmissions, which do not reflect on the quality of care. The focus on all-cause readmissions, rather than condition-specific readmissions, is valuable because it broadens the focus from efforts to improve a narrow set of approaches (such as processes that will prevent recurrent stroke) to more encompassing initiatives (such as medication reconciliation, patient education, follow-up care, and communication between inpatient and outpatient providers) aimed at improving the overall care within the hospital and transitions from the hospital setting. The second limitation is that readmissions are attributed to the hospital, but the hospital may have limited control of care after discharge. The presumption is that the hospital can exert influence on its environment and that there is shared responsibility for postdischarge care, an advantage of restricting the assessments of readmission to only 30 days. Third, a readmission measure could provide an undesirable incentive to deny a patient a needed admission by reducing access for patients, which ultimately could result in a worse outcome.

**Time Frame and Exclusions**

A readmission is defined as a readmission within 30 days of the day of discharge from the index ischemic stroke hospitalization to any hospital. Therefore, if a patient is discharged from hospital A and readmitted within 30 days of discharge to hospital B, then this readmission should be assigned to hospital A.

Patients with the following characteristics should be excluded from the calculation of the readmission measure: those who died during the index hospitalization, those who left the index hospitalization against medical advice, patients discharged to hospice, and those who were transferred to another acute care facility (the final discharging facility is the one that would be eligible for the calculation of a readmission measure for patients who are cared for by more than one hospital). Moreover, within a given data set, if a patient has 2 hospitalizations for ischemic stroke within a 30-day period, the first hospitalization should be considered the index event, and the second hospitalization should be considered the readmission.

The approach to exclusions of specific readmissions reflects what has been proposed by CMS. Readmissions that are attributable to an expected and appropriate consequence of the index stroke admission should not be included as a readmission for the purpose of evaluating hospital care quality. For example, planned readmissions for carotid endarterectomy or carotid stenting should not be included in the readmission outcome measure, unless the procedure was performed after a patient presented with a recurrent stroke or other acute condition. In contrast, patients with stroke are often readmitted for pneumonia or other pulmonary reasons, which would be included as a readmission.37,99 Admissions to an inpatient rehabilitation facility after discharge from an index stroke hospitalization are not included as a readmission for the purpose of this quality measure; however, hospitalizations that occur from the inpatient rehabilitation facility back to an acute hospital would be counted as a readmission.

Another potential approach focuses on potentially preventable readmissions,100 which include only readmissions that fall into one of several categories believed to be potentially preventable, such as complications plausibly related to care during admission or continuation of reason for initial admission. All patient refined diagnosis-related groups are used to classify reasons for hospitalizations and the relationship between admission and readmission.101 The use of this method to filter readmissions for inclusion in public reporting calculations requires systematic evaluation.

**Review of Evidence on Models That Predict Readmissions After Ischemic Stroke**

Lichtman et al102 conducted a systematic review of the literature about risk-adjustment models for the prediction of stroke readmissions (Table 3). Although no studies were identified that reported a hospital-level risk-adjustment system, 16 publications were identified that reported risk-adjustment models at the patient level.98,103–117 This systematic review describes specific patient characteristics that have been associated with stroke readmissions at the patient level, including age, longer index hospital length of stay, worse physical functioning after stroke, and increased number of hospitalizations before stroke.118 It would not be appropriate to include length of stay of the index hospitalization in risk-adjustment models built for the purpose of evaluating hospital care quality given that longer lengths of stay may be associated with poorer quality stroke care.118

Many published patient-level readmission risk-adjustment models include a measure of stroke severity.102 Few published data are available to inform the discussion regarding the importance of including a measure of stroke severity in risk-adjustment models used to evaluate outcomes at the hospital level (as opposed to the patient level). In an analysis reported by Fonarow et al17 of 33 102 patients cared for at 1 of 404 hospitals that participate in the AHA/American Stroke Association’s GWTG-Stroke program, 30-day readmission rates were similar before and after adjustment for stroke severity with the NIHSS. The unadjusted 30-day hospital readmission rates were as follows: 10th percentile, 5.9%; 90th percentile, 22.9%. The rates adjusted for NIHSS and other patient characteristics were 5.6% and 18.9%, respectively. Therefore, considerable variation in readmission rates across hospitals was observed both with and without adjustment for NIHSS. However, information about how measures of model performance differed for models that included all
Table 3. Characteristics of Publications Examining Predictors of Readmission After Stroke Hospitalization

<table>
<thead>
<tr>
<th>Study</th>
<th>Data Source and Study Period</th>
<th>Study Location</th>
<th>No. of Hospitals/No. of Patients</th>
<th>Stroke Type</th>
<th>Study Outcome</th>
<th>Follow-up Period</th>
<th>Unadjusted Readmission Rate</th>
<th>Analytical Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bohannon and Lee, 2003</td>
<td>Hospital administrative data, stroke center-specific database</td>
<td>United States (Connecticut)</td>
<td>1/326</td>
<td>Ischemic</td>
<td>Same-hospital all-cause readmission</td>
<td>1 y</td>
<td>32.5%</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Bohannon and Lee, 2004</td>
<td>Hospital administrative data, medical chart abstraction, patient interviews</td>
<td>United States (Connecticut)</td>
<td>1/228</td>
<td>Ischemic</td>
<td>Same-hospital all-cause readmission</td>
<td>1 y</td>
<td>37.3%</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Camberg et al, 1997</td>
<td>VA hospital administrative data, Medicare data</td>
<td>United States</td>
<td>Multiple (No. not specified)/2261</td>
<td>Ischemic or hemorrhagic</td>
<td>All-cause readmission</td>
<td>30 d, 6 mo, 1 y</td>
<td>16.6% (30 d), 43.8% (6 mo), 58.8% (1 y)</td>
<td>Proportional hazards regression</td>
</tr>
<tr>
<td>Chuang et al, 2005</td>
<td>Patient interviews</td>
<td>Taiwan (Taipei)</td>
<td>7/489</td>
<td>Ischemic or hemorrhagic</td>
<td>All-cause readmission</td>
<td>30 d</td>
<td>24.3%</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Heller et al, 2000</td>
<td>Hospital patient administrative system, hospital separation database</td>
<td>Australia (New South Wales)</td>
<td>22/1075</td>
<td>Ischemic or hemorrhagic</td>
<td>Unplanned stroke-related readmission</td>
<td>1 y</td>
<td>13%</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Jia et al, 2007</td>
<td>VA hospital administrative data, Medicare data, Medicaid data</td>
<td>United States (Florida)</td>
<td>Multiple (No. not specified)/1818</td>
<td>Ischemic or hemorrhagic</td>
<td>All-cause readmission, stroke readmission</td>
<td>1 y</td>
<td>62.2%, (all-cause), 31.1% (stroke)</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Johansen et al, 2006</td>
<td>Canadian Health Person-Oriented Information Database</td>
<td>Canada</td>
<td>Multiple (No. not specified)/2448</td>
<td>Ischemic or hemorrhagic</td>
<td>All-cause readmission</td>
<td>1 y</td>
<td>37.1%</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Kane et al, 1998</td>
<td>Patient interviews, medical record abstraction</td>
<td>United States</td>
<td>52/457</td>
<td>Stroke type unspecified</td>
<td>All-cause readmission</td>
<td>1 y</td>
<td>30%</td>
<td>Instrumental variables estimation</td>
</tr>
<tr>
<td>Kennedy, 2005</td>
<td>State database of hospital administrative data</td>
<td>United States (California)</td>
<td>Multiple (No. not specified)/38,468</td>
<td>Ischemic or hemorrhagic</td>
<td>No. of stroke events</td>
<td>1 y</td>
<td>12%</td>
<td>Truncated negative binomial regression</td>
</tr>
<tr>
<td>Lichtman et al, 2009</td>
<td>Medicare data (2002)</td>
<td>United States</td>
<td>5070/366,551</td>
<td>Ischemic</td>
<td>All-cause readmission, stroke-related readmission</td>
<td>30 d</td>
<td>14.5% (all-cause), 7.8% (stroke-related)</td>
<td>Cox proportional hazards regression</td>
</tr>
<tr>
<td>Roe et al, 1996</td>
<td>Medical record abstraction</td>
<td>Australia</td>
<td>1/264</td>
<td>Stroke type unspecified</td>
<td>Same-hospital unplanned all-cause readmission</td>
<td>30 d</td>
<td>6.5%</td>
<td>Log-linear analysis</td>
</tr>
<tr>
<td>Smith et al, 2005</td>
<td>HMO administrative data and Medicare/Medicaid data</td>
<td>United States (primarily eastern half, 93 metropolitan counties)</td>
<td>422/9003</td>
<td>Ischemic</td>
<td>All-cause readmission, stroke readmission</td>
<td>30 d, 1 y (for 30 d survivors)</td>
<td>14.0% (30 d all-cause), 40.6% (1 y all-cause), 9.4% (30 d stroke)</td>
<td>Cox proportional hazards regression</td>
</tr>
<tr>
<td>Smith et al, 2006</td>
<td>HMO administrative data, Medicare data, and Medicaid data</td>
<td>United States (11 metropolitan regions)</td>
<td>Multiple (No. not specified)/44,099</td>
<td>Ischemic</td>
<td>All-cause readmission, stroke readmission</td>
<td>30 d</td>
<td>12.9% (all-cause), 7.4% (stroke)</td>
<td>Cox proportional hazards regression</td>
</tr>
</tbody>
</table>

(Continued)
Table 3. Continued

<table>
<thead>
<tr>
<th>Study</th>
<th>Data Source (Study Period)</th>
<th>Study Location</th>
<th>No. of Hospitals/No. of Patients</th>
<th>Stroke Type</th>
<th>Study Outcome</th>
<th>Follow-up Period</th>
<th>Unadjusted Readmission Rate</th>
<th>Analytical Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun and Toh, 2009†‡§††</td>
<td>National Healthcare Group administrative database (2005–2007)</td>
<td>Singapore</td>
<td>3/6464</td>
<td>Ischemic</td>
<td>All-cause readmission, stroke readmission</td>
<td>1 y</td>
<td>37.4% (all-cause), 10.5% (stroke)</td>
<td>Cox proportional hazards regression</td>
</tr>
<tr>
<td>Tseng and Lin, 2009†††‡†</td>
<td>National Health Insurance Research Database (2000–2001)</td>
<td>Taiwan</td>
<td>Multiple (No. not specified)/468</td>
<td>Ischemic, hemorrhagic, TIA, ill-defined, or late effects</td>
<td>All-cause readmission</td>
<td>1 y</td>
<td>50%</td>
<td>Logistic regression</td>
</tr>
</tbody>
</table>

HMO indicates health maintenance organization; TIA, transient ischemic attack; and VA, Veterans Administration.

*Authors only report risk-adjusted models at 30 d.
†Stroke-related readmission represents a composite outcome that includes stroke. Heller et al, 2000†*: readmission for stroke or other cardiovascular disease; Lichtman et al, 2009††: readmission for stroke or other related complications.
‡Rates calculated from data provided in the article.
§All-cause readmission defined as readmission for any cause other than diabetes mellitus and diabetes-related complications.

Reproduced with permission from Lichtman et al.© 2010 American Heart Association, Inc.

risk-adjustment variables with versus without the NIHSS were not provided.

**Minimum Standards for Model Discrimination and Calibration for Models of 30-Day Readmission for Evaluation of Hospital Quality of Care**

Few studies have explicitly reported the performance characteristics of models to predict short-term readmission after stroke at the patient level. Model discrimination for 30-day readmission reported by Fonarow et al.2 was modest (C statistic=0.59). By comparison, model discrimination for 30-day mortality was significantly better (C statistic=0.70). Model calibration was not reported. Current models used in the public assessment of hospital quality for 30-day readmission after discharge for conditions such as myocardial infarction and heart failure have similar levels of discrimination (C statistic=0.63 for readmission after myocardial infarction discharge, and C statistic=0.60 for readmission after heart failure discharge).19,120

Although the literature is scant, it is likely that future models predicting 30-day readmission used for the assessment of hospital quality will also have marginal discriminatory ability. This relatively poor model performance reflects the reality that readmission risk is driven by multiple factors, many of which are unknown, not measured, or occur only after discharge. Some of these factors are likely related to intrinsic patient characteristics that affect readmission, whereas others may include hospital processes that reflect quality of care.

**Conclusions**

Conclusions regarding the minimum requirements for risk-adjustment data when measuring stroke readmissions for the evaluation of stroke care quality at the hospital level are as follows:

1. Data on the predictors of 30-day readmission for evaluation of quality at a hospital level are limited. On the basis of the review of available data, which involved primarily studies of models designed for patient-level prediction, the inclusion of the following minimum list of variables is probably indicated: age, sex, stroke severity, comorbid conditions, and risk factors. Assessment for significance in models that evaluate performance at the hospital level should be performed.
2. As discussed previously, stroke severity appears to be a significant predictor in hospital-level performance of mortality after stroke, although no data are available for hospital-level performance of 30-day readmissions after ischemic stroke. Using the same rationale as for the outcome mortality, inclusion in the prediction model is recommended.
3. Risk-adjustment models of readmission after stroke that do not include stroke severity or other recommended variables must provide comparable classification of hospital performance as a model that includes stroke severity or other missing variables.
4. Stroke severity and other variables that are used in hospital-level risk-adjustment models of readmissions after ischemic stroke should be standardized across sites so that their reliability and accuracy are equivalent.

Suggestions for future research and applications in models of 30-day readmissions are as follows:

1. We conclude that to improve the utility of risk-adjusted models, future researchers should seek to identify factors that (1) are not typically measured in administrative data sources, (2) predict hospital readmission, (3) are not related to hospital quality, and (4) are unevenly distributed among hospitals. These factors may be important opportunities to further improve risk adjustment for 30-day readmissions.
2. Future work should be directed toward understanding how patient preferences are related to poststroke readmission (eg, preferences favoring discharge to home with visiting nurse support versus discharge to an inpatient rehabilitation facility versus discharge to a skilled nursing facility versus preference for hospice care), whether data about patient preferences should be included in models, and if so, how these data can be included in administrative data.
Thirty-Day Functional Status

Strengths and Weaknesses of Functional Status as a Measure of Hospital-Level Quality of Care in Ischemic Stroke

There is both support for and challenges with adopting functional status as a measure of the quality of acute care for patients with ischemic stroke. Global disability-adjusted life years lost because of stroke are projected to grow from 38 million in 1990 to 61 million in 2020.22 In the United States, stroke is a leading cause of major long-term disability among adults. In fact, patients are much more likely to be left with disability than to die of their stroke. An estimated 2 million stroke survivors in the United States currently cope with permanent stroke-related disabilities,4 and between 25% and 74% require some assistance or are fully dependent on caregivers for activities of daily living.121 Functional status after stroke is a central issue to patients and their families and has a powerful impact on their lives. Assessment of functional outcomes, therefore, has very high face validity and is of high relevance to multiple groups, including almost all patients and employers, health insurers, the local community, and healthcare providers. Importantly, strong evidence exists for the impact of care on functional outcomes.10,64,69,122 Most therapies for acute stroke are aimed at minimizing impairment after stroke and improving functional outcomes, and only through the measurement of functional outcomes can institutions that do a better job of providing these therapies be recognized for their accomplishments. In addition, functional measures have much greater sensitivity for the detection of differences between patient groups or changes within patient groups than patients’ vital status.

The greatest challenge with the measurement of functional status is the ability to assess it reliably and systematically across stroke patients in different settings. First, in clinical practice, functional status is currently not uniformly measured in the United States across different healthcare settings. There is no standard functional status assessment for patients in acute care. The Functional Independence Measure (FIM) is used in inpatient rehabilitation facilities and acute hospital rehabilitation units, the Outcome and Assessment Information Set (OASIS) is used in home health care, and the Minimum Data Set (MDS) is used in skilled nursing facilities. Each uses different scales and assesses different bodily functions, activities, and areas of involvement.123 These setting-specific measures and other instruments designed to measure a more basic level of function are also limited by either floor or ceiling effects.124,125 To improve the sensitivity to detect differences between populations provided by these scales and to provide a systematic method for obtaining this information, the CMS is working to implement an instrument that can be used across different healthcare settings.126 A second limitation is the potential costs associated with the adoption of functional status as an indicator of acute healthcare quality, particularly for assessment of postdischarge functioning. Because patients with ischemic stroke have variable patterns of care after the acute hospitalization,127 the introduction of a standard measure at a specific time after the event creates opportunities to measure and improve care, but it also introduces challenges, including locating the patient and obtaining a valid proxy response if it is not possible to obtain information directly from the patient. A final limitation is that 30-day functional status may provide an incomplete insight into a patient’s ultimate recovery, because patient functioning can improve after 30 days. Moreover, this recovery is not solely under the locus of control of the acute hospital, because the care provided after acute discharge, including rehabilitative care, can improve outcomes. Approximately 50% to 70% of stroke patients in the United States are discharged from acute care to receive some form of postacute rehabilitative or nursing care,128,129 so the assessment of functional status after hospital discharge is likely to include the influence of factors not associated with the acute inpatient stay. In other countries where hospital lengths of stay are longer, rehabilitation may be included in the initial hospitalization.130 These challenges are great but not insurmountable relative to the importance of measuring an outcome meaningful to patients, families, health systems, and society.

Definitions of Functional Status After Ischemic Stroke

For the purposes of this report, we adopted the definitions of the World Health Organization’s International Classification of Functioning, Disability and Health.11,24 Functioning is an umbrella term for physiological functions of body systems (eg, neuromusculoskeletal, sensory, mental functions), activities (“execution of a task or action”), and participation (“involvement in a life situation”). Disability is an umbrella term for problems, loss or significant deviation in body functions (eg, visual, cognitive, speech impairments), activity limitations, and participation restrictions (problems experienced in involvement).11 In adopting a more global definition of function that includes both functioning and disability, we were able to examine a broader selection of objective and subjective measures for their applicability as indicators of quality of acute stroke care.

Review of Evidence on Models That Evaluate Predictors of Functional Outcomes After Ischemic Stroke

Of the 706 articles reviewed that related to functional outcomes after ischemic stroke, 19 studies that included 23 analytic models were identified as most relevant to explaining functional outcomes after ischemic stroke (Table 4).131–149 Among the 19 studies that examined functional status at 30 or 90 days, none were designed to evaluate the quality of care at the hospital level or predict outcomes at the provider level. Each study examined functional status outcomes at the patient level, with 13 studies specifying an independent variable. Of those, 7 examined the relationship of specific patient characteristics (including history of diabetes mellitus, serum glucose levels on admission, pretreatment systolic blood pressure, pretreatment Alberta Stroke Program Early CT score [ASPECTS], pretreatment antiplatelet therapy, white matter hyperintensities, gene variants, brachial artery flow-mediated dilation, and posterior circulation infarction), and 6 examined hospital interventions (thrombolytic therapy, tissue-type plasminogen activator–induced recanalization, intravenous tissue-type plasminogen activator, and acute statin therapy).

Studies were conducted in 18 countries with sample sizes ranging from 120 to 4390 (Table 3). The NIHSS was the most
Table 4. Characteristics of Studies That Examined Functional Status

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Location</th>
<th>No. of Patients/No. Analyzed</th>
<th>Stroke Severity Measure</th>
<th>Outcome (Instrument)</th>
<th>Favorable (Poor) Outcome</th>
<th>Timing (Actual)</th>
<th>Data Source for Outcome Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denti et al, 2010131</td>
<td>Italy</td>
<td>1555/1549</td>
<td>GCS, OCSP, SSS, loss of trunk control, indwelling catheter</td>
<td>Rankin Scale</td>
<td>mRS 0–2 (3–5)</td>
<td>30 d</td>
<td>In-person in clinic or telephone interview with patient or caregiver</td>
</tr>
<tr>
<td>Di Carlo et al, 2006132</td>
<td>France, Germany, Hungary, Italy, Portugal, Spain</td>
<td>2472/1983</td>
<td>NR</td>
<td>Rankin Scale</td>
<td>mRS 0–1 (2–5)</td>
<td>90 d</td>
<td>In-person interview with patient or proxy</td>
</tr>
<tr>
<td>Finocchi et al, 1996133</td>
<td>Italy</td>
<td>351/308</td>
<td>GCS, limb paresis classification</td>
<td>Oxford Disability Scale</td>
<td>ODS 0–2 (3–5)</td>
<td>30 d</td>
<td>NR</td>
</tr>
<tr>
<td>German Stroke Study Collaboration, 2004134</td>
<td>Germany</td>
<td>1470/1357</td>
<td>NIHSS</td>
<td>Barthel Index</td>
<td>BI 95–100 (&lt;90)</td>
<td>90 d</td>
<td>Telephone interview</td>
</tr>
<tr>
<td>Kaarisalo et al, 2005135</td>
<td>Finland</td>
<td>4390/3616</td>
<td>NR</td>
<td>Categorical description</td>
<td>Totally or partially independent (completely dependent)</td>
<td>30 d</td>
<td>Medical record</td>
</tr>
<tr>
<td>Lin et al, 2011136</td>
<td>China</td>
<td>2683/2166</td>
<td>NIHSS</td>
<td>Barthel Index</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>Telephone interview or letter questionnaire</td>
</tr>
<tr>
<td>Liou et al, 2010137</td>
<td>Taiwan</td>
<td>187/187</td>
<td>NIHSS</td>
<td>Barthel Index, Rankin Scale</td>
<td>BI 95–100 (&lt;90), mRS 0–1 (2–5)</td>
<td>30 d</td>
<td>In-person clinic or telephone for severely disabled patients</td>
</tr>
<tr>
<td>Maguire et al, 2011138</td>
<td>Australia</td>
<td>640/520</td>
<td>NIHSS</td>
<td>Barthel Index, Rankin Scale, Glasgow Outcome Scale</td>
<td>BI 95–100 (&lt;90), mRS 0–2 (3–5), GOS 1–2 (3–5)</td>
<td>90 d</td>
<td>Telephone follow-up “primarily”</td>
</tr>
<tr>
<td>McCarron et al, 2000139</td>
<td>Scotland</td>
<td>189/152</td>
<td>NIHSS, OCSP</td>
<td>Barthel Index, Rankin Scale</td>
<td>BI &gt;55 (&lt;55), mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>“Observed prospectively”</td>
</tr>
<tr>
<td>Ni Chroinin et al, 2011140</td>
<td>Ireland</td>
<td>448/441</td>
<td>NIHSS</td>
<td>Barthel Index</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>In-person or telephone interview by trained staff, medical record as required</td>
</tr>
<tr>
<td>Ntaios et al, 2010131</td>
<td>Switzerland</td>
<td>1446/NR</td>
<td>NIHSS</td>
<td>Barthel Index</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>In-person in clinic by Rankin-certified personnel</td>
</tr>
<tr>
<td>Rundek et al, 2004142</td>
<td>Canada, United States</td>
<td>1604/1264</td>
<td>NIHSS</td>
<td>Barthel Index</td>
<td>BI 95–100 (&lt;90)</td>
<td>90 d</td>
<td>NR</td>
</tr>
<tr>
<td>Santos-Garcia et al, 2009143</td>
<td>Spain</td>
<td>120/120</td>
<td>NIHSS</td>
<td>Rankin Scale</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>NR</td>
</tr>
<tr>
<td>Sharma et al, 2010144</td>
<td>Singapore</td>
<td>130/114</td>
<td>NIHSS</td>
<td>Rankin Scale</td>
<td>mRS 0–1 (2–5)</td>
<td>90 d</td>
<td>NR</td>
</tr>
<tr>
<td>Tao et al, 2010145</td>
<td>China</td>
<td>1962/1797</td>
<td>NIHSS</td>
<td>Rankin Scale</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>NR</td>
</tr>
<tr>
<td>Tsivgoulis et al, 2007146</td>
<td>Canada, Spain, United States</td>
<td>351/292</td>
<td>NIHSS</td>
<td>Rankin Scale</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>Medical records</td>
</tr>
<tr>
<td>Tsivgoulis et al, 2008147</td>
<td>Canada, Spain, United States</td>
<td>192/160</td>
<td>NIHSS</td>
<td>Rankin Scale</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>Medical records</td>
</tr>
<tr>
<td>Uttenboogaart et al, 2008148</td>
<td>The Netherlands</td>
<td>301/301</td>
<td>NIHSS</td>
<td>Rankin Scale</td>
<td>mRS 0–2 (3–5)</td>
<td>90 d</td>
<td>Medical records</td>
</tr>
<tr>
<td>Weimar et al, 2002149</td>
<td>Germany</td>
<td>4264/2458</td>
<td>NIHSS</td>
<td>Barthel Index, Rankin Scale</td>
<td>BI 95–100 (&lt;90), mRS 0–1 (2–5)</td>
<td>90 d</td>
<td>Telephone interview by trained interviewers or mail questionnaire if not reached</td>
</tr>
</tbody>
</table>

BI indicates Barthel Index; GCS, Glasgow Coma Scale; GOS, Glasgow Outcome Scale; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; NR, not reported; OCSP, Oxford Community Stroke Project; ODS, Oxford Disability Scale; and SSS, Scandinavian Stroke Scale.

common measure of stroke severity (78.9% of studies) at presentation. Functional status at 30 and 90 days was measured with the Barthel Index (BI), Rankin Scale (abbreviated as mRS because not all studies indicated that the modified version was used, even when it was used), Oxford Disability Scale (known as Oxford or Oxfordshire Handicap Scale), and
an investigator-derived 3-level categorical measure. The BI is a cumulative score in multiples of 5, with a range of 0 for completely dependent to 100 for independence in 10 basic activities of daily living with varying weights.101 The mRS is used by clinicians to quantify 6 levels of functional recovery (from a score of 0 [independent] to 5 for patients who require constant care). The Oxfordshire Handicap Scale is an adaptation of the mRS used to quantify “handicap” in 6 levels (score of 0–5) as it relates to an individual’s lifestyle and ability to live independently.150

There are no existing recommendations for how to best analyze the entire range of data collected with any one of these scales. A review of stroke outcome measures from 2000 found most ordinal scales were dichotomized, and there was little agreement in what was considered a favorable or positive outcome.6 Among the 4 measures of functional status in this review, a favorable outcome was analyzed 6 different ways at 2 time points: 5 different ways at 30 days (BI 95–100, mRS 0–1, mRS 0–2, Oxfordshire Handicap Scale 0–2, and “totally or partially independent”) and 4 different ways at 90 days (mRS 0–1, mRS 0–2, BI 60–100, and BI 95–100). The mRS was used as an outcome assessment in 15 of the 23 analytic models reviewed. Five studies did not report the source of information for functional status as an outcome; of those that did, 4 reported the mRS, BI, and investigator-derived scale came from medical records; 6 gathered information with patients in person; and 7 obtained the information by telephone. Only 3 studies specified whether the patient was the only source of information or whether a proxy provided information to evaluate function as the outcome.

The construction of the risk models varied substantially across studies. All 19 studies categorized the outcome into 2 levels and used logistic regression to analyze predictors of function after stroke. Four studies presented 5 analytical models for functional status at 30 days (Table 5), and 15 studies presented 18 analytical models of functional status at 90 days (Tables 6 and 7). There were 82 unique variables analyzed, and only 8 were examined in >50% of the models (age, sex, history of atrial fibrillation, history of diabetes mellitus, history of hypertension, current smoker, prestroke level of disability, and stroke severity). At 30 days, at least half of the models examined age, sex, history of diabetes mellitus, and stroke severity by NIHSS or the Scandinavian Stroke Scale. Findings were similar for studies that examined function at 90 days, with >50% of the models including age; sex; history of diabetes mellitus, atrial fibrillation, and hypertension; a prestroke mRS; and the NIHSS. Every study examined medical history or presence of comorbid conditions, but there was no standard set of conditions. Three studies (16%) included a process measure (onset to treatment time), and none of the models included hospital characteristics or examined differences in outcomes across hospitals.

We were unable to identify a study that evaluated 30- or 90-day functional status outcomes for ischemic stroke patients and the relationship with acute stroke care at the provider or hospital level.

**Choice of Functional Status Measure**

The BI and mRS were the most commonly used scales in the literature. In contrast to the BI, which is widely used in acute trials, the mRS is more responsive to change, has not demonstrated ceiling effects,149 and provides for assessment of the outcome measure of participation. Variable interrater reliability is a potential weakness of the mRS that can be improved through the use of structured interviewing and scoring of raters that is blinded to the treatment provided.151,152

Our synthesis of the literature supports findings from previous reviews that the most common approach is to dichotomize the mRS for analysis. However, there is not a consistent cut point for such dichotomization, where some use ≤1 and other use ≤2 as indicating a favorable outcome. The interpretation of the former is complete independence and the latter is no significant disability. At this time, the lack of a universally accepted level for dichotomization is problematic for recommending an outcome for public reporting on hospitals’ quality of stroke care.

A study of poststroke recovery by Duncan and colleagues153 highlights the challenges in arbitrarily setting a time point for assessment and a cutoff for recovery using the mRS. They showed that assessment of function using the mRS within the first 14 days of the stroke identified 1.74% of the sample as recovered when analyzed as mRS=0 to 1, and 12.2% were recovered when an mRS of 0 to 2 was used. At 30 days, 5.5% and 21.6%, respectively, were recovered, and by 6 months, 24.4% and 54.8% had attained mRS=0 to 1 and mRS=0 to 2 levels of function.

A greater concern is that by dichotomizing scales for the assessment of hospital quality, it is possible that important differences in care may be obscured. For example, if 2 hospitals were to treat an identical patient and the patient in the first hospital achieved an mRS of 5 (requires constant care) and the patient in the second hospital achieved an mRS score of 3 (needs assistance with instrumental but not basic activities of daily living), this difference would be obscured by either dichotomization scheme. Recently, several experts have examined approaches that use the range of data, such as ordinal or sliding dichotomous, depending on the study design and sample size.44,45 These methods, however, are not yet widely used and may be sensitive to the distribution of specific data sets, and the results may be more difficult to interpret.45,154

**Conclusions**

Conclusions regarding the instrument and measurement of functional status at 30 days for the evaluation of stroke care quality at the hospital level are as follows:

1. With full acknowledgment of the limitations of making a firm recommendation, the use of the mRS as a global measure of functional status, administered by a trained mRS rater, may be considered to assess function at 30 days.

2. Ideally, dichotomization of outcomes should be avoided; however, until additional research can help determine the best approach, the assessment of hospital quality with an mRS score of 0 to 2 at 30 days as a conservative estimate of “good” outcome may be considered, with the knowledge that many patients continue to improve over time.

Conclusions regarding the minimum requirements for risk-adjustment data when measuring functional status at 30 days...
for the evaluation of stroke care quality at the hospital level are as follows:

1. The following minimum list of variables should be included in risk-adjustment models: age, sex, stroke severity, comorbid conditions and risk factors (including prior stroke or transient ischemic attack, prior myocardial infarction, coronary artery disease, atrial fibrillation, diabetes mellitus, hypertension, hyperlipidemia, smoking, and alcohol use), prestroke physical function (most commonly measured with the mRS), stroke severity, and stroke type. Assessment for significance in models that evaluate performance at the hospital level should be performed.

2. Premorbid functioning and stroke severity are not collected in a uniform fashion in patients with stroke. They appear, however, to be significant predictors of functional outcomes at 30 days in patient-level performance of mortality after stroke, although no data are available for hospital-level performance of functional outcomes after...
### Table 6. Significance of Variables Associated With Functional Status as Measured by the Rankin Scale at 90 Days

<table>
<thead>
<tr>
<th>Candidate Variable</th>
<th>mRS 0–1</th>
<th>mRS 0–2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociodemographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>NR</td>
<td>NS</td>
</tr>
<tr>
<td>Sex</td>
<td>NR</td>
<td>NS</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Living alone</td>
<td>...</td>
<td>S</td>
</tr>
<tr>
<td>Comorbidities/risk factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior stroke</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Prior TIA</td>
<td>NS</td>
<td>S</td>
</tr>
<tr>
<td>Prior MI</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>CHD or CAD</td>
<td>...</td>
<td>NS</td>
</tr>
<tr>
<td>Valvular HD</td>
<td>...</td>
<td>NS</td>
</tr>
<tr>
<td>Other CVD</td>
<td>...</td>
<td>NS</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Hypertension</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Renal dysfunction</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Smoking</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Alcohol dysfunction</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>BMI</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Family history</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Prestroke status</td>
<td>S</td>
<td>...</td>
</tr>
<tr>
<td>Preadmission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statin therapy</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Antiplatelet therapy</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Antihypertensive</td>
<td>...</td>
<td>NS</td>
</tr>
<tr>
<td>Stroke severity</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Stroke origin, type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCSP, Trial of Org subtypes, PCI/ACI</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Cardioembolic stroke</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Symptomatic ICH</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Occlusion site</td>
<td>...</td>
<td>NS</td>
</tr>
<tr>
<td>Significant arterial pathology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIBI score</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Complete recanalization</td>
<td>...</td>
<td>S</td>
</tr>
<tr>
<td>Stroke imaging characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASPECTS score or infarct volume</td>
<td>...</td>
<td>S</td>
</tr>
</tbody>
</table>

Downloaded from http://stroke.ahajournals.org/ by guest on January 1, 2018
stroke. Using the same rationale as for the outcome mortality, inclusion in the prediction model is recommended.

3. Risk-adjustment models of functional status after ischemic stroke that do not include stroke severity or other recommended variables must provide comparable classification of hospital performance as a model that includes stroke severity or other missing variables.

4. Stroke severity and other variables that are used in hospital-level risk-adjustment models of functional status after ischemic stroke should be standardized across sites so that their reliability and accuracy are equivalent.

The following are suggestions for future research related to measuring functional status after stroke at the hospital level, in order of priority:

1. Functional status is not uniformly captured, clinically or administratively, after stroke. Implementation research...
is urgently needed to examine the feasibility of assessing functional status at a predetermined time (30 days) for every stroke patient and to determine the resources and process for establishing a standardized measure of functional status (including training of assessors) that can be administered reliably.
2. Research is needed to explore the variation in functional status across hospitals and the contribution of specific hospital characteristics after adjustment for patient characteristics, including stroke severity at presentation.

3. Rehabilitation services after stroke have been clearly shown to improve functional status. The receipt of rehabilitation may depend on patient health insurance status, as well as functional status at time of hospital discharge; patients with severe poststroke impairment or minimal impairment are less likely to receive rehabilitation services. Research is needed to determine how best to adjust for receipt of these services after discharge when hospital-based performance is evaluated.

4. Research is needed to guide the field on how to best analyze the mRS as an outcome when hospital quality is examined.

Summary

Each of the outcome measures evaluated in the present scientific statement has strengths and weaknesses. Impairment in functional status is the most common sequel of stroke and is of high relevance to patients, providers, and society. The feasibility of systematic assessment of functional status after stroke, however, is unclear. Information on mortality can be obtained from administrative data sources and has been used for public reporting for several other disease processes. Short-term stroke mortality, however, may reflect patient and family preferences, which currently are not quantifiable nor included in any models. High-quality patient-centered care that allows patients to avoid burdensome life-sustaining therapies may result in a hospital with a higher mortality rate than a hospital that ignores patient preferences and institutes aggressive care for all patients. Readmission is a measure that focuses on healthcare use and likely has less impact in the long term for patients than functional status or mortality.

Each of these 3 measures has the potential to provide information on different aspects of quality of care at the hospital level. Functional status measures may reflect stroke interventions and stroke-related processes; stroke mortality reflects a delicate balance between evidence-based and preference-based care; and the readmission measure may best reflect discharge coordination processes. An examination of a combination of these measures may provide the best overall evaluation of hospital care, with measurement of functional status having the most direct application for measuring the quality of stroke-specific care.

The public reporting of stroke quality information is presumed to motivate quality improvement by targeting information to patients, referring physicians, and payers to select higher-quality hospitals and to incentivize hospitals and physicians to compete on quality by identifying areas for performance improvement. It is critical to note that the link between quality of hospital care and outcomes has not been demonstrated robustly. In addition, the use of case-mix–adjusted outcomes may not reliably reflect differences in quality of care. Unintended consequences, in addition to inaccurate labeling of hospitals as good or poor performers, potentially include incentives for hospitals and physicians to avoid sicker patients, discount patient preferences to achieve outcome goals, and divert resources away from equally or more important quality improvement projects. In addition, there could be a negative effect on provider morale when profiling systems have significant misclassification errors. Finally, of relevance to stroke mortality, another potential negative or unintended consequence is the possibility of steering patients and families away from a more palliative approach in the setting of clear evidence to avoid prolonging the dying process. Balancing the potential for benefit against these risks is no easy task and will require close scrutiny and ongoing monitoring.

Although we focused our review on 3 main outcome measures (mortality, readmission, and functional status), there are multiple additional measures that may prove to be useful when evaluating quality of hospital care, such as complications and length of stay. These measures are easily obtainable but are potentially subject to gaming by the manipulation of hospital coding or the altering of discharge practices.

An important goal of the present scientific statement was to make recommendations on a minimal set of variables to be used in risk adjustment of models that evaluate care at the hospital level. There is a striking lack of data on prediction models for stroke outcomes at the hospital level. It is therefore difficult to make definitive statements regarding the importance of specific variables in risk-adjustment models for the evaluation of hospital performance. The paucity of available evidence to identify a minimal data set underscores the importance and urgency of additional research to support quality assessment.

Variables that are common in all models include age and comorbid conditions, as well as general vascular risk factors. Severity of stroke, defined as either severity of impairment or functional status, has been thought to be universally important. It has strong face and content validity and has been shown to be important for individual patient prediction. For this and all other variables, it is theoretically possible that despite its importance in individual patient prediction models, it may not provide significant information in models that evaluate hospital-level performance if the distribution is similar across hospitals. However, hospitals that have comprehensive stroke programs may receive patients with more severe strokes through direct transport, which would preclude them from exclusion in public reporting programs. With the development of systems of care for stroke in the United States and certification of comprehensive stroke centers, the direct transport by local emergency medical services in addition to the subsequent transfer of patients with severe strokes to selected stroke centers is likely to increase. Importantly, some data are now available that demonstrate a significant impact of adjustment for stroke severity on hospital performance for 30-day mortality, which supports clinical impressions of its importance in risk-adjustment models in the evaluation of performance at the hospital level.

The healthcare system is progressively moving toward outcomes-based care, which will be used to drive reimbursement and compel healthcare systems to focus on and optimize the measurable outcomes of care. Advances in technology will allow clinical information to be extracted more easily from medical record systems and may aid in the assessment
of outcomes after hospital discharge. Along with the opportunity to advance the care of patients with stroke, we must vigilantly seek to minimize the unintended consequences of measuring hospital-based performance by use of inadequate risk adjustment.

Future Research
There is a pressing need for research in multiple areas to better identify methods and metrics to evaluate outcomes of stroke care. In addition to the focused research needs described within each outcome measure section, the following areas are highlighted in order of priority:

1. There is a profound lack of data demonstrating the impact of quality of hospital care on stroke outcomes. It is critically important to define modifiable aspects of stroke care that can improve outcomes.
2. Along similar lines, there is a pressing need to determine whether current case-mix adjustment methods can adequately discriminate quality of care provided by hospitals.
3. Defining the factors, including stroke severity, that are most important to include in risk adjustment of hospital-level models of mortality, readmissions, and functional outcomes 30 days after a stroke is a critical research priority.
4. Research is urgently needed to evaluate the ability of models without the recommended list of variables (particularly stroke severity) to produce comparable discrimination in hospital performance as models that are otherwise similar but include these variables.
5. Future work should be directed toward understanding the roles of depression, family functioning, and caregiver support as intervening variables between acute care and poststroke functional status, mortality, and readmissions need to be explored.
6. Future research is warranted to understand what variables after “time zero” should be included in risk-adjustment models for outcomes after stroke.
7. Research is needed to identify the predictive ability of risk-adjustment models based on automated data collection without manual chart review. These models could include information derived from administrative billing codes or electronic health records.
8. Research is needed on the distribution of patients with varying stroke severity across hospitals and on the referral patterns for patients with severe stroke.
9. Research is needed to identify the most appropriate way to deal with patients transferred in or out of hospitals. Because of the current lack of data, these high-risk patients are typically excluded from determinations of hospital risk-adjusted outcomes. These patients will represent an increasingly important subgroup as systems of care for stroke develop in the United States and transfers increase.
10. The combination of multiple end points, such as dependence or disability and death, is common in randomized controlled trials as a method to gain statistical power; however, the inclusion of such a combination as the primary outcome measure reduces the interpretability of the outcome. Research is needed to evaluate the use of composite end points as outcome measures.
11. The difficulties in coalescing the data from literature reviewed for the present scientific statement highlight the need for greater clarity in reporting of prediction models. Clinical, health services, and outcomes research need to more clearly report how data were collected, the variable structure, and how all variables performed in a model.
12. More research is needed to understand the impact of using in-hospital measures as proxies of the more desirable but often impractical longer-term outcomes.
## Disclosures

### Writing Group Disclosures

<table>
<thead>
<tr>
<th>Writing Group Member</th>
<th>Employment</th>
<th>Research Grant</th>
<th>Other Research Support</th>
<th>Speakers’ Bureau/ Honoraria</th>
<th>Expert Witness</th>
<th>Ownership Interest</th>
<th>Consultant/ Advisory Board</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irene L. Katzan</td>
<td>Cleveland Clinic</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>John Spertus</td>
<td>Saint Luke’s Hospital of Kansas City; University of Missouri–Kansas City</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>Genentech*</td>
<td>None</td>
</tr>
<tr>
<td>Janet Prvu Bettger</td>
<td>Duke University</td>
<td>AHRQ†; PCOR†</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Dawn M. Bravata</td>
<td>Department of Veterans Affairs</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Cheryl Bushnell</td>
<td>Wake Forest University Health Sciences</td>
<td>NIH/NINDS†</td>
<td>World Federation of Neurology*</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Randall T. Higashida</td>
<td>UCSF</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Judith A. Hinchey</td>
<td>Caritas Steward St. Elizabeth’s Medical Center</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Robert G. Holloway</td>
<td>University of Rochester Medical Center</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>George Howard</td>
<td>University of Alabama at Birmingham</td>
<td>Abbott*</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Rosemarie B. King</td>
<td>Northwestern University</td>
<td>NIH†</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Harlan M. Krumholz</td>
<td>Yale University</td>
<td>Medtronic†</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>United Healthcare†; Centers for Medicare &amp; Medicaid Services (CMS, contract to develop performance measures)†</td>
<td>None</td>
</tr>
<tr>
<td>Barbara J. Lutz</td>
<td>University of North Carolina-Wilmington</td>
<td>HRSA†; NIH/NINR†; Rehabilitation Nursing Foundation grant (mentor); VA HSR&amp;D grant*</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>General Electric (stock ownership)†</td>
<td>American Case Management Association (unpaid consultant)†</td>
<td>None</td>
</tr>
<tr>
<td>Matthew J. Reeves</td>
<td>Michigan State University</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Eric E. Smith</td>
<td>University of Calgary</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>Genentech*</td>
<td>None</td>
</tr>
<tr>
<td>Robert W. Yeh</td>
<td>Massachusetts General Hospital</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

This table represents the relationships of writing group members that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all members of the writing group are required to complete and submit. A relationship is considered to be “significant” if (1) the person receives $10,000 or more during any 12-month period, or 5% or more of the person’s gross income; or (2) the person owns 5% or more of the voting stock or share of the entity, or owns $10,000 or more of the fair market value of the entity. A relationship is considered to be “modest” if it is less than “significant” under the preceding definition.

*Modest.
†Significant.
### Reviewer Disclosures

<table>
<thead>
<tr>
<th>Reviewer Disclosures</th>
<th>Employment</th>
<th>Research Grant</th>
<th>Other Research Support</th>
<th>Speakers’ Bureau/ Honoraria</th>
<th>Expert Witness</th>
<th>Ownership Interest</th>
<th>Consultant/ Advisory Board</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mohammed</td>
<td>University of Birmingham, UK</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Mohammed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patrick Romano</td>
<td>University of California Davis</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Gustavo Saponsnik</td>
<td>St Michael's Hospital, University of Toronto</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

This table represents the relationships of reviewers that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all reviewers are required to complete and submit. A relationship is considered to be "significant" if (1) the person receives $10,000 or more during any 12-month period, or 5% or more of the person’s gross income; or (2) the person owns 5% or more of the voting stock or share of the entity, or owns $10,000 or more of the fair market value of the entity. A relationship is considered to be "modest" if it is less than "significant" under the preceding definition.

### References


5. Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. BMC Health Serv Res. 2007;7:91.


30. The COPSS-CMS White Paper Committee; Ash AS, Fienberg SE, Louis TA, Normand S-LT, Stukel TA, Utts J. Statistical Issues in Assessing...
Stroke  March 2014


113. Roe CJ, Kulinskaya E, Brisbane M, Brown R, Barter C. A methodol-
ogy for measuring clinical outcomes in an acute care teaching hospital. 

114. Smith MA, Frytak JR, Liou JJ, Finch MD. Rehospitalization and sur-
vival for stroke patients in managed care and traditional Medicare plans. 

115. Smith MA, Liou JJ, Frytak JR, Finch MD. 30-Day survival and re-
hospitalization for stroke patients according to physician specialty. 

116. Sun Y, Toh MP. Impact of diabetes mellitus (DM) on the health-care uti-
lization and clinical outcomes of patients with stroke in Singapore. *Value 

117. Tseng MC, Lin HJ. Readmission after hospitalization for stroke in Taiwan: results from a national sample. *J Neurol 


JD, Stauffer BD, Bernhard SM, Epstein AJ, Wang Y, Herrin J, Chen J, 
Federer JJ, Mattera JA, Krumholz HM. An administrative claims me-
sure suitable for profiling hospital performance on the basis of 30-day 

120. Krumholz HM, Lin Z, Drye EE, Desai MM, Han LF, Rupp MT, Mattera 
JA, Normand SL. An administrative claims measure suitable for profiling 
hospital performance based on 30-day all-cause readmission rates among 

121. Kalra L, Langhorne P. Facilitating recovery: evidence for organized 

122. Dennis MS, Lewis SC, Warlow C, FOOD Trial Collaboration. Routine 
oral nutritional supplementation for stroke patients in hospital (FOOD): a 

123. Prvu Bettger JA, Caster WI, Latham NK, Keysor JJ. Analyzing change in 
recovery patterns in the year after acute hospitalization. *Arch Phys Med 

124. Dromerick AW, Edwards DF, Diringer MN. Sensitivity to changes in dis-
bility after stroke: a comparison of four scales useful in clinical trials. 

125. Kwon S, Hartzema AG, Duncan PW, Min-Lui S. Disability mea-
sures in stroke: relationship among the Barthel Index, the Functional 

126. Accredited Disease Payment Reform Demonstration Center for 
MedicareDemonstrationProjects/DemoProjectsEvalRpt/Medicare 


129. Dennis MS, Lewis SC, Warlow C, FOOD Trial Collaboration. Routine 

130. Prvu Bettger JA, Caster WI, Latham NK, Keysor JJ. Analyzing change in 
recovery patterns in the year after acute hospitalization. *Arch Phys Med 

131. Dromerick AW, Edwards DF, Diringer MN. Sensitivity to changes in dis-
bility after stroke: a comparison of four scales useful in clinical trials. 

132. Kwon S, Hartzema AG, Duncan PW, Min-Lui S. Disability mea-
sures in stroke: relationship among the Barthel Index, the Functional 

133. Accredited Disease Payment Reform Demonstration Center for 
MedicareDemonstrationProjects/DemoProjectsEvalRpt/Medicare 


136. Dennis MS, Lewis SC, Warlow C, FOOD Trial Collaboration. Routine 

137. Prvu Bettger JA, Caster WI, Latham NK, Keysor JJ. Analyzing change in 
recovery patterns in the year after acute hospitalization. *Arch Phys Med 

138. Dromerick AW, Edwards DF, Diringer MN. Sensitivity to changes in dis-
bility after stroke: a comparison of four scales useful in clinical trials. 

139. Kwon S, Hartzema AG, Duncan PW, Min-Lui S. Disability mea-
sures in stroke: relationship among the Barthel Index, the Functional 

140. Accredited Disease Payment Reform Demonstration Center for 
MedicareDemonstrationProjects/DemoProjectsEvalRpt/Medicare 


Risk Adjustment of Ischemic Stroke Outcomes for Comparing Hospital Performance: A Statement for Healthcare Professionals From the American Heart Association/American Stroke Association

Irene L. Katzan, John Spertus, Janet Prvu Bettger, Dawn M. Bravata, Mathew J. Reeves, Eric E. Smith, Cheryl Bushnell, Randall T. Higashida, Judith A. Hinchey, Robert G. Holloway, George Howard, Rosemarie B. King, Harlan M. Krumholz, Barbara J. Lutz and Robert W. Yeh

Stroke. published online January 23, 2014;

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/early/2014/01/23/01.str.0000441948.35804.77

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/