Public Reporting of Quality Data for Stroke
Is It Measuring Quality?

Adam Kelly, MD; Joel P. Thompson, MPH; Deborah Tuttle, RN, MPS; Curtis Benesch, MD, MPH; Robert G. Holloway, MD, MPH

Background and Purpose—Public reporting of quality data is becoming more common and increasingly used to improve choices of patients, providers, and payers. We reviewed the scope and content of stroke data being reported to the public and how well it captures the quality of stroke care.

Methods—We performed a cross-sectional survey of all report cards within the Agency for Healthcare Research and Quality Report Card Compendium. Stroke quality data were categorized into one of 5 groups: structure, process, outcomes, utilization, and finances. We also determined the congruence of mortality ratings of New York hospitals provided by 2 different report cards.

Results—Of 221 available report cards, 19 (9%) reported quality information regarding stroke and 17 specifically addressed the quality of hospital-based stroke care. The most frequent data reported were utilization measures (n=15 report cards) and outcome measures (n=14 report cards). Data regarding finances (n=4), structure of care (n=2), and process of care (n=1) were reported infrequently. Ratings were incongruent in 61 of the 157 hospitals (39%) with the same hospital being rated below average on one report care and average on another in 44 hospitals.

Conclusions—Publicly reported quality data pertaining to patients with stroke are incomplete, confusing, and inaccurate. Without further improvements and a better understanding of the needs and limitations of the many stakeholders, targeted transparency policies for stroke care may lead to worse quality and large economic losses. (Stroke. 2008;39:3367-3371.)

Key Words: mortality ■ quality of care ■ stroke

P
ublic reporting of healthcare quality data has become a common policy strategy to improve transparency, accountability, and quality. It is hoped that the power of information will increase trust and drive better choices by patients, referring physicians, and purchasers of health care. Public reporting of hospital quality data is incorporated into federal law and many states have a mandatory public reporting requirement.1,2 Transparency policies, however, can decrease quality and lead to large economic losses if the information provided is incomplete, confusing, inaccurate, or distorted.3,4

Stroke care is of high priority on the national quality agenda. Outcomes measures are readily available (using administrative data), process measures have been developed from research evidence, and the Joint Commission and state health departments have encouraged the adoption of primary stroke centers.5–11 Quality initiatives have also shown that improvement on some of the measures is achievable.12,13 Despite these observations, the scope and accuracy of publicly reported quality data regarding stroke care are unknown.

In November 2006, the Agency for Healthcare Quality and Research (AHRQ) released its Report Card Compendium, which assembled all publicly available web sites reporting health quality data at a single Internet site.14 We reviewed the individual report cards to assess the content and scope of publicly available information regarding quality data after stroke. We also investigated the degree to which different sites provided similar estimates of quality regarding the same hospital.

Materials and Methods
All report cards in the AHRQ Report Card Compendium were accessed from a central summary web page.14 From this central web page, links were followed and each web site was reviewed for the presence of stroke-specific quality data. For each report card that contained stroke information, we identified the sponsor(s), the geographic scope of the quality report, the number of hospitals included in the report card, and the timeliness of the data. In addition, we identified the data sources used, how patients with stroke were defined, and the methods of risk adjustment, if used.15

Each quality measure was categorized as a measure of structure, process, outcome, utilization, or finance.16,17 Structural measures include the availability and quality of resources, management systems, and designation as a primary stroke center. Process measures (performance measure or effectiveness measures) evaluate the activities of physicians and other healthcare providers to determine if evidence-based recommendations are followed. Outcome measures evaluate the end result of health care and include measures of mortality, readmissions, complications, and patient/caregiver satis-
fication scores. Utilization measures include data pertaining to the frequency of service use, including length of stay. Finance measures include economic data pertaining to the provision of stroke care.

We identified the rating of all New York state hospitals as provided by 2 separate report card sites: the Niagara Health Quality Coalition (NHQC), a not-for-profit corporation, and Health Grades, a for-profit, nationwide evaluation program. Both sites report on inpatient stroke mortality, use International Classification of Diseases, 9th Revision, Clinical Modification diagnosis codes to define their population at risk, and exclude patients transferred to another acute care hospital. Minor differences in coding rules exist between the 2 reporting systems (eg, Health Grades includes International Classification of Diseases, 9th Revision code 436 and excludes patients with a palliative care code, V66.7). Neither report card provided ratings of hospitals with low stroke volumes, which was defined as <30 cases per year. New York state was chosen as the source for this analysis due to the availability of 2 report cards, which separately evaluated nearly all hospitals statewide.

The NHQC accepts no advertising, consulting, or other funding from the providers it grades and uses statewide hospital administrative data (age 18 years and older) for the calendar year 2005. NHQC uses software developed by the AHRQ to compares each hospital’s inpatient mortality with a risk-adjusted national average (AHRQ Inpatient Quality Indicators). This measure is risk-adjusted using a linear model estimated from a nationwide data set and includes age, gender, and All-Payer Refinements. The groups as developed by 3M. NHQC uses a 95% CI to identify which hospitals are better (3-star rating), worse (1-star rating), or not significantly different from the statewide average (2-star rating).

Health Grades uses Medicare inpatient billing data (age 65 years and older) for the period 2003 to 2006 and uses a proprietary “disease-specific and outcome-specific” risk-adjustment methodology using demographic characteristics, comorbid diagnoses, and specific procedures. Health Grades calculates a predicted mortality for each hospital using a proprietary risk-adjusted methodology, to which the actual hospital mortality is compared. The statistical methodology determining if the actual and predicted rates are significantly different is not readily available on the web site. Those hospitals whose mortality was lower than predicted were assigned 5-star ratings and those hospitals whose mortality was higher than predicted were assigned a one-star rating. Those hospitals whose actual performance was not significantly different from what was predicted received a 3-star rating (2-star and 4-star ratings are not assigned). According to methodology found on the company’s web site, for each diagnosis or procedure, approximately 70% to 80% of actual performance was not significantly different from what was predicted.

For each hospital in New York state that was included in both report cards, we assessed agreement between the 2 reporting systems for the following categorical ratings: (1) not significantly different than the state average or from what was predicted (ie, 2-star rating from NHQC and 3-star rating from Health Grades); (2) better than the state average or lower than predicted (ie, 3-star rating from NHQC and 5-star rating from Health Grades); and (3) worse than the state average or higher than predicted (ie, one-star rating from NHQC and one-star rating from Health Grades). We performed similar analyses for hospitals designated by the Joint Commission as primary stroke centers and hospitals designated by New York state as designated stroke centers. Weighted kappa was used to assess level of agreement.

This study was exempt from review by the University of Rochester’s Research Subject Review Board.

Results
A total of 221 online quality reports were included in the AHRQ compendium as of December 1, 2007. From this total, 5 could not be found from the central AHRQ compendium site and 11 were proprietary for or health plan members only. From the remaining 205 report cards, 19 (9%) reported quality information regarding stroke care. Of these 19 report cards, 17 were hospital-based. One provided physician self-reported measures regarding stroke prevention care, and one indicated the presence of stroke management services for health plans. Of the 205 accessible reports, 76 contained data regarding hospital-based measures. Thus, 17 of 76 hospital quality reports (24%) included stroke-specific quality measures.

The 17 hospital-based report cards are listed in Table 1. The report cards were produced by state health departments, independent research organizations or private–public partnerships, insurance companies, for-profit companies, hospital associations, and health systems. All sites used hospital administrative data that were 1 to 3 years old, and 13 sites used the Inpatient Quality Indicator risk-adjustment software provided free by the AHRQ. One site created separate reports for nonhemorrhagic and hemorrhagic strokes. All other sites grouped subarachnoid hemorrhages, intracerebral hemorrhages, and ischemic strokes into one measurement cohort. Two sites included supplemental survey information about the structure and process of care.

Table 2 shows the type of quality measures reported. The most frequently reported measures were utilization (n=15) and outcome (n=14) data. Risk-adjusted inpatient stroke mortality was the most commonly reported outcome measure. Two sites included quality indicators about the structure of stroke care services and only one site (from the United Kingdom) addressed the process of stroke care. Four sites included economic information with various methods of presentation (eg, costs, average charges, median charges).

The results of comparing inpatient mortality ratings for similar hospitals are summarized in Table 3. A total of 157 of 214 New York state hospitals were evaluated by both the NHQC and Health Grades (the majority of the 57 unrated hospitals were due to low volume). Health Grades rated 56 of the 157 hospitals below average compared with only 16 rated below average by NHQC. The 2 sites provided congruent ratings (average/average, below average/below average, above average/above average) on 96 of these 157 hospitals (agreement 61.1%, weighted kappa 0.163, slight agreement above that expected by chance). Ratings were incongruent in 61 of 157 cases (38.9%), including one case in which a hospital was rated above average by one site and below average by the other. The most common disagreement was an average rating by NHQC and a below average rating by Health Grades (44 of the 157 hospitals [28%]). Only one hospital was rated as above average by both sites. Agreement was similarly low for the 10 primary stroke centers (agreement 40%, weighted kappa 0.118) and the 107 of the 116 designated stroke centers evaluated by both sites (agreement 61.6%, weighted kappa 0.173).

Discussion
The science of quality measurement is maturing at a rapid and frenetic pace. In evaluating healthcare delivery, good quality is no longer assumed. On the contrary, there is an increasing expectation that it should be measured, compared, and paid for if good results are to be achieved. Our study provides a window into what is currently being publicly reported regarding the quality of stroke care, although the results of our study are limited to only those report cards included in the AHRQ
Many more organizations, health systems, and hospitals are also likely reporting stroke quality data on the Internet. The amount and content of data available in countries outside the United States remain uncertain. The results are concerning for several reasons. First, the data are incomplete. Despite there being well-established process measures for stroke, they were reported by only one site and this involved hospitals from the United Kingdom. Few sites reported on the structural elements of quality (stroke unit, accredited facility, designated stroke center).

### Table 1. Summary of Hospital-Based Report Cards Including Stroke Quality Data

<table>
<thead>
<tr>
<th>Sponsor(s)</th>
<th>Title of Report Card</th>
<th>Geographic Area</th>
<th>No. of Hospitals</th>
<th>Year of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>About.com, Inc.</td>
<td>UCompare Health Care</td>
<td>Nationwide (US)</td>
<td>5500+</td>
<td>2005</td>
</tr>
<tr>
<td>Blue Cross Blue Shield of Minnesota</td>
<td>Healthcare Facts</td>
<td>Multistate (Minn, eastern ND, western Wis)</td>
<td>33</td>
<td>2007</td>
</tr>
<tr>
<td>Blue Cross Blue Shield of Tennessee</td>
<td>Hospital Quality Comparison</td>
<td>Statewide (Tenn)</td>
<td>77</td>
<td>2004–2005</td>
</tr>
<tr>
<td>Colorado Health and Hospital Association Performance and Quality Group</td>
<td>Colorado Hospital Quality</td>
<td>Statewide (Colo)</td>
<td>65</td>
<td>2004–2006</td>
</tr>
<tr>
<td>Dr Foster Ltd</td>
<td>Hospital Guide</td>
<td>Nationwide (UK)</td>
<td>167</td>
<td>2006</td>
</tr>
<tr>
<td>Florida Agency for Health Care Administration</td>
<td>Florida Compare Care</td>
<td>Statewide (Fla)</td>
<td>206</td>
<td>2006</td>
</tr>
<tr>
<td>Fraser Institute</td>
<td>Ontario Hospital Report Card</td>
<td>Provinciwide (Ontario)</td>
<td>136</td>
<td>2005</td>
</tr>
<tr>
<td>Health Grades</td>
<td>Free Hospital Rankings</td>
<td>Nationwide (US)</td>
<td>4776</td>
<td>2004–2006</td>
</tr>
<tr>
<td>Kentucky Hospital Association</td>
<td>Quality Data</td>
<td>Statewide (Ky)</td>
<td>66</td>
<td>2006</td>
</tr>
<tr>
<td>Maryland Health Care Commission</td>
<td>Hospital Guide</td>
<td>Statewide (Md)</td>
<td>47</td>
<td>2006</td>
</tr>
<tr>
<td>Massachusetts Executive Office of Health and Human Services</td>
<td>Healthcare Quality and Cost Information</td>
<td>Statewide (Mass)</td>
<td>71</td>
<td>2004–2005</td>
</tr>
<tr>
<td>Niagara Health Quality Coalition; Alliance for Quality Health Care</td>
<td>2007 New York State Hospital Report Card</td>
<td>Statewide (NY)</td>
<td>231</td>
<td>2005</td>
</tr>
<tr>
<td>Norton Healthcare</td>
<td>Quality Report</td>
<td>Multi-state (Ind, Ky)</td>
<td>4</td>
<td>2006–2007</td>
</tr>
<tr>
<td>Office for Oregon Health Policy &amp; Research</td>
<td>Oregon Hospital Quality Indicators</td>
<td>Statewide (Ore)</td>
<td>57</td>
<td>2005</td>
</tr>
<tr>
<td>PacifiCare</td>
<td>Hospital Performance</td>
<td>Multistate (Calif, Wash, Ore, Texas, Okla, Nev, Ariz)</td>
<td>200</td>
<td>2004</td>
</tr>
<tr>
<td>Texas Department of State Health Services</td>
<td>Indicators of Inpatient Care in Texas Hospitals</td>
<td>Statewide (Texas)</td>
<td>287</td>
<td>2004</td>
</tr>
</tbody>
</table>

Report Card Compendium. Many more organizations, health systems, and hospitals are also likely reporting stroke quality data on the Internet. The amount and content of data available in countries outside the United States remain uncertain. The results are concerning for several reasons.

### Table 2. Specific Content of Quality Data Included in Report Cards

<table>
<thead>
<tr>
<th>Structure (2 sites)</th>
<th>Process (one site)</th>
<th>Outcomes (14 sites)</th>
<th>Utilization (15 sites)</th>
<th>Financial (4 sites)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of a stroke unit</td>
<td>No. of admissions given CT scans within 24–48 hours</td>
<td>Risk-adjusted in-hospital mortality rate</td>
<td>Case volume</td>
<td>Average charge</td>
</tr>
<tr>
<td>No. of beds in stroke unit</td>
<td>Arrangement for continued physiotherapy on discharge</td>
<td>Risk-adjusted complication rate</td>
<td>Risk-adjusted average charge</td>
<td>Cost</td>
</tr>
<tr>
<td>Presence of senior stroke clinician in hospital</td>
<td>No. of patients taking aspirin on admission who were given a nonaspirin antplatelet at discharge</td>
<td>No. of in-hospital deaths</td>
<td>Risk-adjusted length of stay</td>
<td>Median charge</td>
</tr>
<tr>
<td>Integration of care with community health and social service organizations</td>
<td>Review of stroke cases by a neurologist</td>
<td>Risk-adjusted in-hospital + 30-day mortality rate</td>
<td>Average length of stay</td>
<td>Median charge</td>
</tr>
<tr>
<td>No. of patients nursed in acute stroke unit</td>
<td>Risk-adjusted in-hospital + 180-day mortality rate</td>
<td>Risk-adjusted 15-day readmission rate</td>
<td>30-day readmission rate</td>
<td>30-day readmission rate for complications and/or infection</td>
</tr>
<tr>
<td>Designation as a primary stroke center</td>
<td>Review of stroke cases by a neurologist</td>
<td>Risk-adjusted in-hospital + 30-day mortality rate</td>
<td>Average length of stay</td>
<td>Median charge</td>
</tr>
</tbody>
</table>
which does not generally include process or structure data. Quality reporting is the availability of administrative data, patient-centeredness (e.g., patient satisfaction) or health disparities. The majority of stroke deaths occur after discharge from a hospital, and the risk-adjusted in-hospital mortality rate is the most commonly reported outcome measure. This mortality rate is generally lower than the actual risk of death for any specific patient, and it corresponds poorly with process measures of patient care. Short-term mortality correlates poorly with process measures and is likely related to unsafe care in fewer than 10% of all deaths. In fact, the majority of stroke deaths occur after deliberate decisions by patients and their families not to pursue unwanted life-prolonging treatments. Short-term mortality, therefore, may be more indicative of “good quality” deaths, particularly because more informed patients are more inclined to want less aggressive care (i.e., better quality decision-making leading to higher short-term mortality). The tremendous variability in how mortality and other outcome data are reported only compounds the confusion. It is also unclear if the average user knows how to interpret and use other measures that are frequently reported such as utilization data (e.g., length of stay) or financial information (e.g., charges versus costs).

Second, the data are unreliable. We found that 2 separate report cards provided disparate hospital ratings in 39% of comparisons. Disagreement was also observed among primary and designated stroke centers, a subset of hospitals selected for the capacity and quality of stroke care they provide. A recent study showed inconsistent ratings of hospitals among several sites for surgical procedures, but did not quantify the degree of disagreement. The tremendous variability in how mortality and other outcome data are reported only compounds the confusion. It is also unclear if the average user knows how to interpret and use other measures that are frequently reported such as utilization data (e.g., length of stay) or financial information (e.g., charges versus costs).

Third, the data are unreliable. We found that 2 separate report cards provided disparate hospital ratings in 39% of comparisons. Disagreement was also observed among primary and designated stroke centers, a subset of hospitals selected for the capacity and quality of stroke care they provide. A recent study showed inconsistent ratings of hospitals among several sites for surgical procedures, but did not quantify the degree of disagreement. The tremendous variability in how mortality and other outcome data are reported only compounds the confusion. It is also unclear if the average user knows how to interpret and use other measures that are frequently reported such as utilization data (e.g., length of stay) or financial information (e.g., charges versus costs).

Unreliable and invalid publicly reported stroke quality data may have unintended consequences. Patients may choose the wrong providers, payers may reward or punish providers inappropriately, providers may “game” to improve rankings, hospital leaders may divert resources from worthy improvement efforts, and intermediary companies may profit by stoking fears of losing reputation and market share among affected hospitals. In the end, the public loses trust.

We provide 3 recommendations. First, efforts are needed to develop a standardized “dossier” of stroke quality measures that meaningfully align with the 6 worthy aims of health care: effective, safe, patient-centered, equitable, timely, and efficient. This objective will include efforts to harmonize existing stroke process measures (which are in progress) and to develop consensus metrics for stroke outcomes that measure “good-quality” deaths as well as unexpected “never ever” deaths for which organizations should be held accountable. In addition, we need to develop and standardize new measures that focus on patient-centered, efficient, and equitable care. Collaborative public–private partnerships with several organizations that are currently committed to providing stroke quality data for internal quality improvement efforts could facilitate such efforts.

Second, there should be more organized skepticism focused on the AHRQ stroke inpatient quality indicator as a primary measure of quality of care. The increasing appetite for healthcare quality data and the easy access of administrative data will likely guarantee the continued use of mortality as a marker of quality. In the short run, this will placate stakeholders. Fundamental questions remain, however, about the appropriateness of combining all types of stroke (subarachnoid hemorrhage, intracerebral hemorrhage, ischemic) into this one indicator and the impact that such measures may have on the delivery of high-quality palliative care. The inpatient time horizon is confounded by hospital practice patterns and the capacity of nonhospital services and ignores the longitudinal accountability needed to improve the quality of a chronic condition. Finally, despite its “public access,” the risk-adjusted methodology remains proprietary.

Third, further national efforts are needed to develop standardized reporting requirements with explicit rules to reduce bias and to ensure a minimum standard for measuring and reporting conduct. Much can be learned from the transparency systems that help govern corporate financing, restaurant hygiene, and mortgage-lending practices. As the quality field continues to mature, there will be increasing efforts to cherry-pick measures for marketing purposes. All measures should be reported, good or bad; there is no substitute to playing by the rules and working with integrity. Discussion is also needed regarding mandatory versus voluntary reporting, internal reporting with feedback versus public reporting, and how to finance a sustainable and effective transparency system that is responsive, interactive, and customized to stakeholder preferences and public concerns.

Summary

The modest amount of stroke quality data that is currently available does not portray an accurate measure of the quality of care being provided, and inconsistencies in these data further undermine its effective utilization. The future of stroke quality measurement and reporting is uncertain, but broad improvement in the science and infrastructure is needed to realize its potential to mobilize choices and market forces to improve stroke care. Providers must take a leading role in these efforts and focus on the needs of our patients and the public at large.
Disclosures

R.G.H. consults for Milliman Guidelines reviewing practice guidelines and Maximus, Inc.

References


43. Holloway RG, Quill TE. Mortality as a measure of quality; implications for palliative and end-of-life care. JAMA. 2007;298:802–804.


