Towards a National System for Monitoring the Quality of Hospital-Based Stroke Services

Nicolas Weir, MB ChB, MRCP(UK); Martin S. Dennis, MB ChB, FRCP; on behalf of the Scottish Stroke Outcomes Study Group

Background and Purpose—We sought to evaluate a system for monitoring the quality of hospital-based stroke services that uses routinely collected case fatality data, adjusted for case mix, as well as simple measures of the process of stroke care.

Methods—We compared the process of care and case fatality after stroke between 5 Scottish hospitals (A through E) during 1995–1997. We retrospectively identified 2724 patients with acute stroke using routine discharge data and obtained case mix and process of care data from the medical record. We ascertained case fatality by record linkage and adjusted for case mix using a simple, externally validated regression model.

Results—Crude case fatality varied by 21 deaths per 100 admissions between the 5 hospitals. After adjustment, case fatality still differed significantly (P=0.047), with 5 to 7 more deaths per 100 admissions at Hospital A than at Hospitals B through E. There were major shortcomings in the specialization and organization of care, the use of CT scanning, and the completeness of documentation at Hospital A compared with the other hospitals. There were smaller, but clinically important, differences in care between Hospitals B through E but no significant differences in adjusted case fatality.

Conclusions—Once adjusted for important prognostic variables, routinely collected case fatality data might identify hospitals with major shortcomings in the processes of stroke care. More moderate, but still clinically important, variations in stroke care can only be identified by monitoring the process of care directly. (Stroke. 2001;32:1415-1421.)

Key Words: case fatality ■ cerebrovascular disorders ■ outcome ■ stroke management

In recent years a considerable degree of public and professional concern has been expressed about variations in the standards of healthcare provision.1–5 Fueled by a rash of medical scandals, the rise of consumerism, a political drive to improve accountability, and a genuine desire to focus on quality rather than costs, there have been calls for systems to be put in place to routinely monitor and ensure the quality of clinical services.6–8 Several such mechanisms have recently been introduced to the United Kingdom, including the publication of indicators that aim to describe performance in certain key areas of healthcare.9 Currently, these performance indicators rely on comparisons of routinely collected outcome data to identify hospitals in which the quality of care might be poor. As the third most common cause of death and the leading cause of disability among adults in the United Kingdom, the outcome of patients with stroke has been an early candidate for such scrutiny: comparisons of 30-day case fatality and the proportion of patients discharged home after stroke have been published for Scottish hospitals since 1994. Similar discharge data were first published in England and Wales in 1999.90,11

For comparisons of outcome to be fair, however, it is important to compare similar entities. Unless variation in the characteristics of patients admitted (case mix) is taken into account, differences in outcome between hospitals are more likely to reflect differences in referral patterns, admission thresholds, and the nature of the local population than differences in the process of care or its quality.12,13 However, hospital stroke outcomes published in the United Kingdom take into account only variation in age, sex, and socioeconomic circumstance (Scotland) or age and hospital type (England and Wales) and therefore fail to adjust for more important prognostic variables such as prestroke functional status and the severity of the stroke itself. In their current format, therefore, these data may provide misleading indications of the quality of stroke care. Furthermore, they give no information on the process of care and therefore cannot be used to target the aspects of care that require improvement. Whether collection of some simple measures of process and more complete adjustment for case mix would improve our ability to differentiate between hospitals with different standards of stroke care is unknown. The aim of this study was to investigate this possibility and to test a system that might practicably be used to routinely monitor the quality of hospital-based stroke services.

Received December 11, 2000; final revision received March 2, 2001; accepted March 5, 2001.
A list of the members of the Scottish Stroke Outcomes Study Group appears in the Appendix.
Correspondence to Dr Martin Dennis, Department of Clinical Neurosciences, Western General Hospital, Crewe Rd, Edinburgh EH4 2XU, Scotland.
E-mail msd@skull.dcn.ed.ac.uk
© 2001 American Heart Association, Inc.
Stroke is available at http://www.strokeaha.org

1415
Subjects and Methods

Hospitals and Subjects
We retrospectively studied stroke patients admitted to 5 Scottish hospitals (A through E) between August 1995 and July 1997. We selected the hospitals to span the range of the first “league table” of stroke outcomes published in Scotland in 1994.10 Two were university teaching hospitals, and 3 were district general hospitals. We used routine hospital discharge data to identify all patients discharged with an International Classification of Diseases (ICD) cerebrovascular disease code listed as the principal diagnosis. A trained and experienced nurse auditor then inspected the medical records to identify those patients who had had an acute stroke. We accepted this diagnosis if it was the diagnosis assigned by the most senior physician whose opinion was recorded, if onset was 30 days before admission, and if the patient had not been transferred from a hospital that was not part of the study. We determined stroke pathological subtype from the result of a CT head scan or autopsy. We excluded patients with a subarachnoid hemorrhage (SAH). As the study progressed, we stopped collecting ICD codes with a very low positive predictive value for acute, non-SAH stroke: after 6 months, we stopped collecting cases with International Classification of Diseases, Ninth Revision (ICD-9) codes 433, 435, and 438; after 1 year, we stopped collecting cases coded as SAH (ICD-9 code 430; International Classification of Diseases, 10th Revision [ICD-10] code I60).

Case Mix and the Process of Care
We collected data describing case mix and the process of care from physician and nursing entries in the medical record (Tables 1 and 2). We derived socioeconomic status from postal code information, using Carstairs scores to assign a deprivation category to each patient (higher scores indicate greater deprivation).14 We collected process of care variables that are either known to influence survival (organized and specialized rehabilitation15 and the long-term use of antithrombotic drugs after ischemic stroke16,17) or held to represent

### TABLE 1. Study Participants and Their Baseline Characteristics (Case Mix) at Each Hospital

<table>
<thead>
<tr>
<th>Hospital</th>
<th>Eligible cases</th>
<th>Notes audited</th>
<th>Acute strokes*</th>
<th>Age, median and IQR, y</th>
<th>Male</th>
<th>Social class†</th>
<th>Before stroke</th>
<th>Previous history of</th>
<th>Stroke type</th>
<th>Stroke onset in hospital</th>
<th>On admission‡</th>
<th>Proportion predicted to be dead at 6 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>740</td>
<td>94</td>
<td>521</td>
<td>75 67–82</td>
<td>49</td>
<td>7</td>
<td>5</td>
<td>15</td>
<td>46</td>
<td>5</td>
<td>44</td>
<td>43</td>
</tr>
<tr>
<td>B</td>
<td>1187</td>
<td>696</td>
<td>746</td>
<td>73 65–80</td>
<td>46</td>
<td>2</td>
<td>2</td>
<td>15</td>
<td>67</td>
<td>9</td>
<td>70</td>
<td>38</td>
</tr>
<tr>
<td>C</td>
<td>717</td>
<td>94</td>
<td>511</td>
<td>75 67–82</td>
<td>50</td>
<td>2</td>
<td>4</td>
<td>15</td>
<td>48</td>
<td>7</td>
<td>70</td>
<td>38</td>
</tr>
<tr>
<td>D</td>
<td>999</td>
<td>700</td>
<td>720</td>
<td>73 53–82</td>
<td>50</td>
<td>2</td>
<td>6</td>
<td>15</td>
<td>74</td>
<td>9</td>
<td>44</td>
<td>37</td>
</tr>
<tr>
<td>E</td>
<td>580</td>
<td>497</td>
<td>349</td>
<td>75 67–82</td>
<td>47</td>
<td>1</td>
<td>2</td>
<td>15</td>
<td>67</td>
<td>6</td>
<td>45</td>
<td>14</td>
</tr>
</tbody>
</table>

IQR indicates interquartile range; ADL, activities of daily living; and GCS, Glasgow Coma Scale.

*Denominator for calculation of subsequent proportions (%).
†Cases with missing data: deprivation category, 15 cases (A to E, n = 516, 745, 517, 545, 386, respectively); diabetes mellitus and ischemic heart disease, 7 cases (A to E, n = 521, 745, 520, 547, 384, respectively); normal GCS eye and motor score, 16 cases (A to E, n = 518, 742, 520, 545, 383, respectively).
‡At onset if stroke occurred in hospital.
§Kruskal-Wallis test.
good practice (the provision and speed of CT scanning and the completeness of the medical record). Our definition of specialized stroke care was based on that used by the Stroke Unit Trialists Collaboration. We defined a “stroke consultant” as a consultant physician who had specific responsibility for managing the hospital’s stroke service and who had received relevant training. We measured completeness of documentation using the Royal College of Physicians Stroke Audit Form (RCPSAF) in a representative sample of cases at each hospital (every nth case, with n varying to ensure similar numbers across hospitals). The RCPSAF consists of 60 criteria that assess the recording of standard items of history, examination, investigation, and management. The overall RCPSAF score represents the proportion of items recorded, excluding those appropriately omitted because of the patient’s clinical condition, eg, failure to record an assessment of the ability to swallow is disregarded if the patient is unconscious.

**Case Fatality and Adjustment for Case Mix**

We derived 6-month case fatality by linkage to national death certificate data. We did not collect functional outcome because there are currently no systems capable of doing so routinely in Scotland. We adjusted for case mix using 2 methods. In the first, we adjusted for age, sex, and social deprivation to mimic the method currently used to adjust the stroke outcome data routinely published in Scotland. In the second, more comprehensive method, we adjusted for the covariates of a recently described model of survival after stroke (the “study model”). This model was developed from data from the Oxfordshire Community Stroke Project and was validated in 2 independent cohorts. The study model covariates were as follows: age; whether the patient lived alone before the stroke and was independent in simple activities of daily living; and, on admission, whether the patient could speak and was orientated in time and place, could lift both arms against gravity, and could walk without the help of another person. In both methods, we fitted a multiple logistic regression model for 6-month case fatality by forcing in the relevant covariates. To determine whether case fatality was significantly different between hospitals after adjustment, we added a hospital term into each regression. For the study model only, we measured model calibration using the Hosmer-Lemeshow goodness-of-fit C statistic and discrimination by the area under the receiver operating characteristic curve.

**Absolute Difference in Case Fatality Between Hospitals**

We calculated w scores to show the absolute difference in 6-month case fatality between hospitals both before and after adjustment. The w score expresses the difference between the observed and predicted number of deaths per 100 patients admitted and is calculated by the formula \( w = \frac{(o-p)n}{100} \), where \( o \) is the observed number of deaths, \( p \) the predicted number of deaths, and \( n \) is the total number of admissions. Differences between hospital w scores indicate the absolute difference in case fatality per 100 patients admitted. For the unadjusted w scores, we derived \( p \) at each hospital by multiplying the number of cases admitted (\( n \)) by the proportion of cases that died in the 5-hospital cohort overall. For the adjusted w scores, we derived \( p \) at each hospital by summing the individual predicted probabilities of death generated by the prognostic model.

**Statistical Analysis**

We compared proportions using the \( \chi^2 \) test and median values using the Kruskal-Wallis test. We calculated 95% CIs for the w scores using the method described by Parry et al. We performed all analyses using SPSS for Windows (7.5) and Epi Info (6.04b).

**Results**

**Subjects**

We identified 4223 hospital admissions and were able to audit the medical records of 4017 (95% overall, ranging from 93% at Hospital E to 98% at Hospital C). Audited and unaudited cases did not differ significantly in terms of age or sex at any hospital or in 6-month case fatality at Hospitals A, B, D, or E. Unaudited cases at Hospital C were significantly more likely than audited cases to be alive at 6 months but represented only 2% (17 cases) of the hospital sample. After inspecting the medical records, we identified 2724 cases of acute, non-SAH stroke, 1007 (37%) of whom had died by 6 months. We were unable to collect all 6 covariates used by the study model for 19 patients (3 at Hospital A, 5 at Hospital B, 9 at Hospital D, and 2 at Hospital E). We retained these cases in the analysis by inserting the most pessimistic value for any missing variables. Analyses with optimistic values gave similar results.

**Baseline Characteristics**

The baseline characteristics of patients differed significantly between the 5 hospitals (Table 1). On the whole, the prevalence of prognostic variables for predicting death was highest at Hospital A and lowest at Hospital D. These differences are summarized in the different proportions of patients predicted to be dead by 6 months: 43% at Hospital A versus 27% at Hospital D. Baseline prognosis was intermediate at the remaining hospitals; the proportions predicted to be dead by 6 months were 40% at Hospital C, 38% at Hospital B, and 37% at Hospital E.

**Crude and Adjusted Case Fatality**

There were large and highly significant differences in the proportion of patients dead by 6 months between hospitals (Table 2). Comparison of the crude w scores (Figure) showed that, by 6 months, approximately 11 more patients were dead...
per 100 admitted to Hospital A than to Hospitals B, C, and E, and that approximately 10 more patients were dead per 100 admitted to Hospitals B, C, and E than to Hospital D.

After adjustment for age, sex, and social deprivation, the absolute differences in case fatality between hospitals remained substantial (Figure), and the hospital term in the regression analysis was still highly significant ($P<0.0005$). After adjustment for the variables in the study model, the differences in case fatality between Hospital D and Hospitals B, C, and E were virtually abolished, and the differences in case fatality between Hospital A and the other 4 hospitals was reduced to between 5 and 7 more deaths per 100 patients admitted. The hospital term in the regression analysis remained marginally significant ($P=0.047$). The study model showed satisfactory calibration (Hosmer-Lemeshow goodness-of-fit C statistic=$14.21$; $df=10$; $P=0.1636$) and discrimination (area under receiver operating characteristic curve=$0.84$).

### The Process of Stroke Care

Except for the use of antithrombotic drugs after ischemic stroke, the provision of the measured items of stroke care was substantially lower at Hospital A than at the other 4 hospitals (Table 2). In particular, Hospital A did not provide specialized and organized stroke care. All other hospitals operated a stroke rehabilitation unit (Hospitals B and D for 24 months, Hospital E for 18 months, and Hospital C for 11 months) and had a stroke consultant throughout the study. None of the hospitals had an acute stroke unit during the study period. Hospital A obtained a CT head scan in only half of its patients with stroke compared with ≥76% elsewhere. As a result, considerably more patients at Hospital A were prescribed antithrombotic drugs without prior imaging to exclude cerebral hemorrhage. On average, the medical notes at Hospital A complied with the lowest proportion of RCPSAF criteria, and a multidisciplinary team meeting was documented in only 10% of cases.

There were less marked and less consistent differences in the provision of the measured items of stroke care between Hospitals B through E (Table 2). Nonetheless, some of these differences were potentially important, particularly the 2-fold difference in the proportion of patients admitted to a stroke rehabilitation unit (SRU) and the near 2.5-fold difference in the proportion of patients discharged from the care of a stroke consultant. Additionally, Hospitals B and C clearly failed to admit patients younger than 65 years to their SRU, and nearly two thirds of patients admitted to the SRU at Hospital E were not in fact cared for by a stroke consultant. Only half of the CT scans at Hospital B were performed within 7 days of admission, and perhaps some small intracerebral bleeds were therefore missed. Only 15% of patients overall and under half of those admitted to a SRU at Hospital E had the results of a CT scan performed.

## Discussion

This study has 3 main findings. First, we have shown that the large differences in crude case fatality between the study hospitals could be largely explained by important differences in case mix and that much of the remaining differences in case fatality could be attributed to chance. In other words, most of the variation in outcome after stroke between the
study hospitals was due to factors outside their control rather than differences in the quality of the care they provided. Adjustment for only age, sex, and social deprivation, the method currently used to adjust stroke outcomes published in Scotland—a more comprehensive method than that used in England and Wales—was only slightly better than no adjustment at all. Clearly, stroke outcomes published in the United Kingdom are inadequately adjusted for case mix and are unlikely to be accurate indicators of the quality of stroke care.

One of our aims, however, was to investigate whether improved adjustment would allow case fatality data to identify hospitals with different standards of stroke care. Our second finding is that this might be the case when variations in stroke care are substantial. Thus, after adjustment for important differences in case mix, case fatality at Hospital A was significantly higher than that at Hospitals B through E and was associated with the failure of Hospital A to provide any organized and specialized rehabilitation and, compared with the other hospitals, a much lower use of CT head scanning and a lower standard of documentation (according to the RCPSAP criteria). The lack of organized stroke care may have directly contributed to the higher case fatality. The absolute difference in case fatality of approximately 5% was similar to that in the systematic review of randomized trials comparing stroke unit care with that on general wards.\textsuperscript{15} The lower proportion of patients scanned and the less detailed documentation, although not directly therapeutic, might plausibly reflect the resources, priority, and attention that were given to patients with stroke, less tangible factors that may nonetheless have an impact on outcome.\textsuperscript{23} However, this conclusion is based on the findings at a single hospital. Furthermore, despite reasonably large numbers and the best possible case mix adjustment, the residual difference in case fatality between hospitals was relatively small and only marginally statistically significant and might also quite plausibly be explained by residual variation in case mix and/or chance.

Our third main finding is that adjusted case fatality data are clearly unable to differentiate between hospitals with moderate differences in stroke care. Thus, we found significant and potentially important differences between Hospitals B, C, D, and E in terms of access to organized stroke care, the degree to which stroke unit care was truly specialized, the documentation of multidisciplinary team meetings, the provision of and delay in CT scanning, the prescription of antithrombotic drugs, and the completeness of the medical record. Despite this, after adjustment for case mix, the 4 hospitals had almost identical 6-month case fatalities. These 4 hospitals are likely to be representative of many hospitals in the United Kingdom that provide average to good stroke services but that nonetheless have the scope to improve in one or more aspects of care. Not only do outcome data fail to identify such opportunities, but, if relied on as the sole measure of quality of care, they run the risk of engendering an attitude of complacency rather than one of critical self-reflection. On the other hand, by measuring simple aspects of the process of care, even moderate opportunities to improve stroke services can be targeted directly, and at many hospitals rather than a few.

In common with many audit projects, we made use of routinely and retrospectively collected data. As a result, it is possible that some of the differences in outcome, case mix, and the process of care that we observed, and hence our conclusions regarding the relationship between outcome and quality of care, may simply reflect variation in documentation and coding practices. For example, in the absence of a stroke unit or stroke consultant, the recording of baseline characteristics at Hospital A may have been systematically less accurate, and hence adjustment for case mix systematically less complete, than at Hospitals B through E. On the other hand, because we excluded patients who had not in fact had an acute stroke, and because the outcomes at each hospital were standardized against their predicted outcomes, any impact of variation between hospitals in the sensitivity and specificity of their routine discharge data is likely to be small. Except for the RCPSAF, our measurements of the process of care did not take full account of variation in patients’ needs. However, after the data set was stratified into tertiles of predicted risk, ie, into groups of patients with roughly similar treatment needs, the findings of our audit of the process of care remained broadly the same (data available on request). Bias may also have resulted because our auditor was not blind to the hospital or, in many cases, to the patient’s outcome, but this reflects normal practice. Our audit did not address all elements of stroke care, such as the identification and treatment of complications and the adequacy of nursing and therapist input, because these cannot be reliably obtained from the medical record.

Previous studies that have investigated the ability of outcome data to indicate the quality of stroke care have reached conflicting conclusions.\textsuperscript{12,24–30} The only large study to show a reliable relationship between outcome and the quality of stroke care related outcomes to the treatment of individuals rather than to the quality of service.\textsuperscript{29} In our study we identified only one hospital with poor quality of care, and that was the hospital with the highest adjusted case fatality. However, the sensitivity and specificity of using case fatality, adjusted by our methods, to identify hospitals that are performing badly still need to be established. Any system with a low specificity (ie, with a high false-positive rate) for detecting poor-quality services has the potential to waste resources on identifying local problems that may not exist and would almost certainly adversely affect staff morale and patient and public confidence. Because it would require the study of very large numbers of hospitals, perhaps the only realistic way of establishing sensitivity and specificity would be to introduce a national system to routinely and prospectively collect case fatality data, prognostic variables to adjust for case mix, and key measures describing the process of care and organization of services. Somewhat discouragingly, a simulation study suggests that even with perfect adjustment for case mix and adequate sample size, a system based on outcome measurement alone would still fail to reliably identify “bad apples.”\textsuperscript{31}

This study suggests that far more useful comparisons of the results and quality of hospital stroke services might be made if routine systems were to collect both case fatality data, appropriately adjusted for case mix, and important indicators.
of the organization and process of stroke care. Our simple and robust methods might be used as a template for such a system. In this study the measures of process of care were collected by an auditor reviewing medical records. However, in Scotland it will soon be possible to determine from routinely collected data the proportion of patients accessing stroke unit care and who were managed by a stroke consultant; the proportion having a CT scan and its timing; and the proportion discharged on antithrombotic drugs. A recent study has also shown that the collection of RCPSAF data describing the completeness of the medical record in samples of stroke patients is feasible across hospitals at a national level.32 Similarly, because our prognostic variables are uncomplicated, clinical, and measured on admission, they lend themselves to prospective collection by junior medical staff (perhaps prompted by an admission pro forma) and therefore might allow a routine system to avoid the cost and inaccuracy of abstracting prognostic data from the medical records. The prospective collection of our prognostic variables might also improve the accuracy with which routine hospital discharge data identify cases of stroke: those cases that have prognostic data appended are probably more likely to have been admitted with an acute stroke and thus to receive an appropriate diagnostic code.

Our findings are important because we have studied large numbers of patients admitted to several hospitals over a 2-year period and because we have used a simple but powerful and externally validated prognostic model to adjust for case mix. This in turn has allowed us to test a system of measuring the quality of stroke care using patient outcomes and process measures that are both credible and routinely practicable. Given the uncertainties and limitations of using outcomes alone to indicate the quality of stroke care, even with state of the art adjustment for case mix, we suggest that routine systems also should measure simple but robust indicators of the organization and process of care (by periodic case note review, prospective data collection, or, perhaps in the future, with the use of automatic methods based on an electronic patient record33). In Scotland we are currently planning a national system to monitor the quality of hospital-based stroke services that incorporates routine measurement of case mix, process, and outcome.

Acknowledgments

This study was supported by grants from the Chief Scientist’s Office and the Clinical Resource and Audit Group of the Scottish Office. Dr Weir was supported by a Wellcome Trust Health Services research training grant. Dr Martin Dennis is the stroke consultant at one of the hospitals included in this study.

Appendix

The Scottish Stroke Outcomes Study Group: Members and Responsibilities

Dr Nicolas Weir, Wellcome Research Training Fellow: manuscript preparation and data collection, management, analysis, and interpretation; Dr Martin Dennis, Principal Investigator: study design, interpretation, and manuscript preparation; Michael McDowall: study design, software provision, data management, security, audit, and preparation; Alison Gunkel: data collection and audit; Charles Warlow, Co-Investigator: study design and manuscript preparation; Christopher Povey: collaborator at the Information and Statistics Department of the National Health Service, Scotland; Marion Livingstone: data management; Pam McLaren: data management; Dr Peter Langhorne: hospital facilitator; Dr Jackie Taylor: hospital facilitator; Dr Peter Murdoch: hospital facilitator; Dr James McCallion: hospital facilitator; Ian Campbell: hospital facilitator.

References

5. Smith R. All changed, changed utterly: British medicine will be transformed by the Bristol case. BMJ. 1998;316:1917–1918.
Towards a National System for Monitoring the Quality of Hospital-Based Stroke Services
Nicolas Weir and Martin S. Dennis
on behalf of the Scottish Stroke Outcomes Study Group

Stroke. 2001;32:1415-1421
doi: 10.1161/01.STR.32.6.1415

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/32/6/1415

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/