Measurement of Motor Recovery After Stroke
Outcome Assessment and Sample Size Requirements

Pamela W. Duncan, PhD, PT; Larry B. Goldstein, MD; David Matchar, MD; George W. Divine, PhD; and John Feussner, MD

Background and Purpose: The purpose of this study was to analyze recovery of motor function in a cohort of patients presenting with an acute occlusion in the carotid distribution. Analysis of recovery patterns is important for estimating patient care needs, establishing therapeutic plans, and estimating sample sizes for clinical intervention trials.

Methods: We prospectively measured the motor deficits of 104 stroke patients over a 6-month period to identify earliest measures that would predict subsequent motor recovery. Motor function was measured with the Fugl-Meyer Assessment. Fifty-four patients were randomly assigned to a training set for model development; 50 patients were assigned to a test set for model validation. In a second analysis, patients were stratified on basis of time and stroke severity. The sample size required to detect a 50% improvement in residual motor function was calculated for each level of impairment and at three points in time.

Results: At baseline the initial Fugl-Meyer motor scores accounted for only half the variance in 6-month motor function ($r^2=0.53, p<0.001$). After 5 days, both the 5-day motor and sensory scores explained 74% of the variance ($p<0.001$). After 30 days, the 30-day motor score explained 86% of the variance ($p<0.001$). Application of these best models to the test set confirmed the results obtained with the training set. Sample-size calculations revealed that as severity and time since stroke increased, sample sizes required to detect a 50% improvement in residual motor deficits decreased.

Conclusions: Most of the variability in motor recovery can be explained by 30 days after stroke. These findings have important implications for clinical practice and research. (Stroke 1992;23:1084–1089)

Key Words • motor activity • prognosis • rehabilitation

Stroke is the third leading cause of death in the United States and a major source of disability. Although stroke mortality is significantly decreasing because of decreased incidence of stroke as well as improved survival rates, the initial period of high mortality is over survival is good, with 50% of stroke patients alive in 7 years. Most of the survivors of stroke do experience some level of neurological recovery, yet 30–60% of stroke survivors are dependent in some aspects of activities of daily living (ADL). Thus, stroke disability continues to be a problem of major proportions.

Previous studies demonstrate that stroke recovery generally begins early, with the fastest improvement occurring over the first month. However, our knowledge of the details of recovery from stroke remains limited. First, many of the natural history studies of stroke have measured ADL rather than specific motor function. This distinction is important because patients may improve in ADL by compensating for the neurological deficit with the uninvolved side while the actual motor deficit remains unchanged. Second, most studies of recovery after stroke have drawn patient samples from rehabilitation facilities. These studies include patients in whom much of the spontaneous improvement has already occurred. Furthermore, they reflect recovery in a selected patient population that introduces bias and limits the general applicability of the results. Lastly, many studies may have limited generalizability because stroke types were heterogeneous or unspecified, comprehensive and well-characterized measures of motor function were not used, and assessments were not always performed at specified time intervals after stroke.

The purpose of the present study was to analyze recovery of motor function in a cohort of patients with an acute occlusion in the carotid distribution. Our goal was to describe the motor recovery over the first 6 months to identify the earliest measures that would predict the level of motor recovery at 6 months. These data can be used to improve the design of intervention trials by providing a framework for early stratification. Within strata we can provide sample size estimates for trials of interventions aimed at improving final motor recovery. In addition, these data can be clinically useful for estimating patient care needs and establishing therapeutic plans.
Design

This study was a part of a larger study designed to investigate the importance of serum glucose levels on stroke outcome. The patients constitute a cohort of patients admitted with a new stroke between January 1987 and October 1989 at Duke University Medical Center, Durham County General Hospital, or the Durham Veterans Affairs Medical Center. All hospital admissions for stroke in Durham County, North Carolina, occur at one of these three hospitals. The study protocol was approved by the Institutional Review Board of each participating hospital.

Inclusion/Exclusion Criteria

During the study period, 999 patients presented with a presumptive diagnosis of atherothrombotic stroke. To be included in the study, patients had to have 1) give informed consent to participate either personally or by proxy, 2) be over 40 years of age, 3) reside within 100 miles of Duke University Medical Center, 4) be hospitalized within 24 hours of onset of neurologic symptoms, 5) have a measurable neurological deficit on admission, 6) have no preexisting stroke deficit, 7) have a neurological deficit persisting more than 24 hours (transient ischemic attacks were excluded), and 8) lack any medical condition from which death was likely within 6 months. The eligibility criteria were relaxed to include patients with a previous minor stroke when any prior neurological deficit would not interfere with measurement of the clinical course of the new deficit. Patients with hemorhagic, vertebrobasilar artery occlusions, or brain stem strokes as determined by computed tomographic (CT) scan were excluded from the study. Patients were excluded if they had risk factors for embolic stroke, including cardiomyopathy, anterior myocardial infarction in the previous 6 months, new atrial fibrillation, or a prosthetic heart valve. Patients with major medical comorbidities (e.g., amputations, severe chronic obstructive pulmonary disease, severe arthritis) that would limit assessment of motor function.

Outcomes

The major outcomes of the study included the 6-month motor score on the Fugl-Meyer Assessment of stroke and the capacity to perform ADL as measured by the Barthel Index. The latter instrument, demonstrated to be a valid and reliable measure of disability, has been widely used in stroke research. The Fugl-Meyer test measures motor recovery, which is not specifically addressed directly by the Barthel Index. The Fugl-Meyer Assessment is a 226-point scoring system that includes range of motion, pain, sensation, motor function of the upper and lower extremities, and balance. The sensory component of the test comprises 24 points and the motor component 100 points. This instrument provides a reliable and valid measure of specific motor function that is also sensitive to change. The assessments of motor impairment and ADL disability were made by a study nurse or physician's assistant trained to use the Fugl-Meyer and Barthel instruments. Training included review of a videotape plus up to five sessions with a physical therapist experienced in using these scales. Assessments were made within 24 hours of admission and at 5, 30, 90, and 180 days after admission. The evaluators obtained motor and ADL data from patients with stroke by interview and physical examination in the hospital and, after discharge, at the patient's home.

At each assessment period, at least 87% of living patients were evaluated. Only 11% of patients who survived a minimum of 180 days after stroke had fewer than three assessments. Patients were not assessed for function if they were too ill or an appointment for the examination could not be scheduled. Patients with missing data on physical function and ADL status did not differ significantly from those with complete data.

Data Analysis

The dependent variable for the primary analysis was the Fugl-Meyer motor score obtained 6 months after stroke. Multiple regression models were used to predict the 6-month motor score using data available at baseline, 5 days, and 30 days after stroke. Fugl-Meyer motor and sensory scores, the product of the motor and sensory scores, and the Barthel Index scores were selected as potential predictors of 6-month motor recovery. Barthel scores were measured at 5 days and at each examination thereafter.

The data set was randomly divided to provide a training data set and a test data set. The training data set was used to develop the models. The best models at baseline, 5 days, and 30 days after stroke were then applied to the test set for validation. The comparability of the two data sets was assessed with the $t$ test for continuous variables and with Student's $t$ test for categorical variables and with Student's $t$ test for continuous variables. The PC/SAS procedure REG (SAS Institute, Cary, N.C.) was used to compare the predictive models. A model was considered best if it had the highest $R^2$ among the models with all variables significant ($p<0.05$).

In a second analysis, the patients were stratified into four groups based on the Fugl-Meyer motor scores at baseline, 5 days, and 30 days after stroke (0–35, severe; 36–55, moderately severe; 56–79, moderate; and >80, mild). Analysis of the 6-month Fugl-Meyer motor scores and the 6-month Barthel Index scores were performed for each level of stroke severity at each time point. The sample size required to detect a 50% improvement in residual motor function after 6 months was calculated for each measure of impairment at the three points in time after stroke.

Results

One hundred forty-six patients were originally enrolled in the Durham County Stroke Study. They were randomly assigned to the test set or training set at the time of analysis. Twenty-three patients died before the 6-month assessment, 17 did not have 6-month motor scores recorded, and two were excluded because baseline measures could not be obtained. The remaining 104 patients, 54 in the training set and 50 in the test set, were available for analysis. There were no differences in mortality or missing 6-month measurements between the two sets. Table 1 gives the demographic characteristics, initial total Fugl-Meyer scores, and initial Fugl-Meyer motor and sensory subscores for both groups.
None of the differences between the training set and test set were statistically significant. As expected, patients who died had more severe initial motor deficits than those who survived (33.4 versus 57.5, \( p = 0.002 \)).

To assess how well and how soon after stroke 6-month motor recovery could be predicted, the potential predictors of Barthel Index scores and Fugl-Meyer motor and sensory scores were analyzed by regression. Regression analysis performed on the training set data revealed that 53.2% of the variance of 6-month Fugl-Meyer motor score could be explained by the baseline Fugl-Meyer motor score alone (Table 2). The best model based on data available at 5 days after stroke included the motor and sensory scores and explained 74.2% of the variance of 6-month Fugl-Meyer score. The best predictive model after 30 days included only the 30-day motor score and explained 86.2% of the variance in motor scores 6 months after stroke who achieved assisted independence in ADL parallel motor recovery patterns measured with Fugl-Meyer Assessment. Graph represents mean ADL scores by each severity stratum at baseline, 5 days, and 30 days.

The most dramatic recovery in motor function occurred over the first 30 days, regardless of the initial severity of the stroke (Figure 1). However, the moderate and most severe stroke patients continued to experience some recovery for 90 days. Nineteen percent of patients classified as severe at baseline improved to moderate deficits, and none had mild deficits after 6 months.

Recovery of ADL as measured with the Barthel Index paralleled the motor recovery patterns (Figure 2). The correlations between the Fugl-Meyer score and the Barthel Index score at 5, 30, 90, and 180 days after stroke were computed. They ranged from \( r = 0.80 \) to 0.91 \( (p < 0.001) \). Table 3 gives the percentages of patients at each severity stratum at baseline, 5 days, and 30 days after stroke who achieved assisted independence in ADL parallel motor recovery patterns measured with Fugl-Meyer Assessment. Graph represents mean ADL scores by each severity stratum at baseline, 5 days, and 30 days.
ADL (Barthel Index score of ≥60)27 and who had complete functional recoveries (Barthel Index score of 100) 6 months after stroke. (A Barthel Index score of >60 was chosen in this study because Granger and colleagues27 have reported that 60 is a pivotal score.) Most patients with severe initial motor deficits never achieved complete recovery in ADL (Table 3). If the severe motor deficit persisted at 30 days, less than 10% of the patients achieved 6-month Barthel scores of 100, and only 56% achieved >60 points on the Barthel ADL Index.

If therapeutic planning is to be effective, information about both the predictability and rate of recovery after stroke are essential. This information is also valuable in designing clinical trials that evaluate efficacy of intervention. The degree and rate of recovery have significant influences on the sample sizes required to demonstrate an effect of treatment. Table 4 gives sample sizes required to show a 50% improvement in the residual motor deficit after 6 months for each severity stratum at baseline, 5 days, and 30 days after stroke.26 The number of patients needed for a clinical intervention trial is influenced by both the length of time after stroke and the severity of the stroke at any time point.

### Discussion

The purpose of this study was to use a well-characterized measure of motor function to evaluate recovery after acute anterior circulation ischemic stroke. In agreement with previous reports,5,7,8-11 spontaneous recovery in this cohort of patients was almost complete by 1 month. The general prognosis for patients in the entire study cohort was good. Eighty-four percent of patients achieved assisted independence in ADL (Barthel Index score of ≥60) after 6 months. Fifty-eight percent of patients had complete recovery of basic ADL (Barthel Index score of 100). This improvement is greater than that reported in other studies. The Framingham study found that 68% of stroke survivors were functionally independent.28 Other nonpopulation based studies found that 43-62% of stroke patients reached independence in ADL.12,29,30 These studies differed from the present investigation in that they included patients with heterogeneous stroke types, followed patients for varying periods, and in most cases selected patients referred for rehabilitative therapy.

The present study focused on recovery of motor function rather than recovery in ADL. However, it is not surprising that these two measures paralleled each other because motor function is a particularly important determinant of physical function and independence in ADL after stroke in humans.31 Yet analysis of our data

### Table 3. Functional Outcomes of Patients Stratified by Severity and Time After Stroke

<table>
<thead>
<tr>
<th>Fugl-Meyer motor score</th>
<th>60 Points</th>
<th>100 Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>66</td>
<td>32</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>92</td>
<td>13</td>
</tr>
<tr>
<td>Moderate</td>
<td>100</td>
<td>22</td>
</tr>
<tr>
<td>Mild</td>
<td>100</td>
<td>37</td>
</tr>
</tbody>
</table>

5-Day status

<table>
<thead>
<tr>
<th>Baseline status</th>
<th>60 Points</th>
<th>100 Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe</td>
<td>65</td>
<td>31</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>Moderate</td>
<td>100</td>
<td>11</td>
</tr>
<tr>
<td>Mild</td>
<td>100</td>
<td>53</td>
</tr>
</tbody>
</table>

30-Day status

<table>
<thead>
<tr>
<th>Baseline status</th>
<th>60 Points</th>
<th>100 Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe</td>
<td>56</td>
<td>25</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>75</td>
<td>4</td>
</tr>
<tr>
<td>Moderate</td>
<td>100</td>
<td>8</td>
</tr>
<tr>
<td>Mild</td>
<td>100</td>
<td>64</td>
</tr>
</tbody>
</table>

### Table 4. Fugl-Meyer Total Motor Score Subgroup Sample Size Statistics

<table>
<thead>
<tr>
<th>Motor Score</th>
<th>n</th>
<th>Mean</th>
<th>Standard Development</th>
<th>Average Remaining Potential Improvement</th>
<th>Sample size required to detect 50% further improvement*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>32</td>
<td>45</td>
<td>35.7</td>
<td>55</td>
<td>28</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>13</td>
<td>81</td>
<td>25.5</td>
<td>19</td>
<td>115</td>
</tr>
<tr>
<td>Moderate</td>
<td>22</td>
<td>86</td>
<td>20.8</td>
<td>14</td>
<td>140</td>
</tr>
<tr>
<td>Mild</td>
<td>37</td>
<td>96</td>
<td>5.4</td>
<td>4</td>
<td>116</td>
</tr>
<tr>
<td>5-Day status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>31</td>
<td>38</td>
<td>30.1</td>
<td>62</td>
<td>16</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>5</td>
<td>81</td>
<td>26.5</td>
<td>19</td>
<td>124</td>
</tr>
<tr>
<td>Moderate</td>
<td>11</td>
<td>90</td>
<td>9.8</td>
<td>10</td>
<td>62</td>
</tr>
<tr>
<td>Mild</td>
<td>53</td>
<td>96</td>
<td>5.9</td>
<td>4</td>
<td>138</td>
</tr>
<tr>
<td>30-Day status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>25</td>
<td>29</td>
<td>24.4</td>
<td>71</td>
<td>9</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>4</td>
<td>46</td>
<td>27.0</td>
<td>54</td>
<td>17</td>
</tr>
<tr>
<td>Moderate</td>
<td>8</td>
<td>81</td>
<td>10.4</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>Mild</td>
<td>64</td>
<td>96</td>
<td>6.1</td>
<td>4</td>
<td>148</td>
</tr>
</tbody>
</table>

*Sample size required in each of two groups to detect a 50% further improvement (of remaining potential improvement) with 80% power, α=0.05, by the two-tailed test.
reveals that patients may score a perfect 100 on the Barthel Index but continue to have the same level of motor impairment. These findings suggest that the level of disability as measured by the Barthel Index is not solely due to motor function in stroke but may be strongly influenced by other variables such as psychosocial supports (T.A. Glass, J.R. Feusner, and D.B. Matchar, unpublished observations). If the purpose of an intervention is to improve motor function, a motor assessment rather than an ADL assessment should be utilized.

Recovery after stroke is influenced by a variety of biological and environmental factors. Patient age, lesion size and location, incontinence, visuospatial defects, prior stroke, prior functional status, and various social characteristics may impact on functional outcome. However, complicated multivariate models do not necessarily improve the accuracy of prognostic estimates. The magnitude of the initial deficit is probably the most important predictor of eventual outcome. We found that the combination of motor and sensory Fugl-Meyer scores 5 days after stroke predicted 74.2% of the variance in the patients' eventual motor function after 6 months. The Fugl-Meyer motor score 30 days after stroke predicted more than 86% of the variance in 6-month motor function.

These findings have important clinical implications. First, the results suggest that most of the variability in 6-month motor recovery can be predicted during acute hospitalization (5 days after stroke). However, the clinician must be aware that some patients may still have a reasonable chance at improving motor function, yet by 1 month after stroke it is not as likely there will be further improvement in motor performance or ADL. Second, the data raise questions about who should receive rehabilitation. The effectiveness of focused rehabilitative programs for stroke patients remains controversial. Deciding which patients should receive what type of rehabilitative therapy is not certain. Because definitive data are not available, information concerning anticipated recovery patterns is useful in discussing prognosis and making plans for further therapy and long-term disposition. In the present study, all patients with mild or moderate motor deficits at baseline, 5 days, and 30 days after stroke were independent in ADL after 6 months (Table 3). In contrast, only 56% of the patients with a severe deficit 30 days after stroke were independent in ADL after 6 months.

Knowledge of the expected patterns of recovery for stroke patients is also necessary for the rational design of interventional trials. In the present analysis, patients were empirically stratified into four groups based on the Fugl-Meyer scores at three time points within the first month after stroke (Table 3). Prediction using the regression models gave large reductions in residual variance at 6 months, as reflected in the high R² values. However, the residual in 6-month variances using the model were very similar to the variances seen within subgroups of patients defined solely by their motor score status at the same prediction points. That is, simply grouping patients by motor score at a given time point is essentially as effective in reducing variability as using that motor score in a regression equation to predict 6-month motor score. Therefore, in calculating sample size, we used only estimates based on grouping patients by motor scores. For these sample-size calculations, 50% improvement in the residual motor deficit after 6 months was considered clinically meaningful. The number of patients in a given severity stratum required to demonstrate a clinically meaningful difference decreases as the time after stroke increases (Table 4). This decrease occurs because most of the spontaneous recovery is completed by 30 days after stroke. Further, at any given time after stroke, the number of patients required to demonstrate a clinically meaningful difference after 6 months decreases as the severity of the stroke increases. This is because of the "ceiling effect." Patients with severe deficits at early time points have greater deficits after 6 months than patients with less severe deficits (Figure 1). A 50% improvement in patients with early mild deficits may only be several points on the Fugl-Meyer scale, whereas a similar percent improvement in patients with initially severe deficits may be as much as 25–30 points. The number of patients required to show a small difference is greater than the number of patients required to demonstrate a large difference. Some of the sample-size calculations should be viewed with caution because of the small number of patients in some groups. However, these data suggest that it is imperative that researchers be aware of the composition of their stroke population before designing clinical trials. For example, if the patients' deficits are mild and acute, as is more likely in community- and hospital-based populations, it will take more subjects to detect an effect of treatment. In contrast, in rehabilitation settings in which patients are more likely severe and subacute, sample-size requirements will be smaller.

References

2. Wolf PA: An overview of the epidemiology of stroke. Stroke 1990;21(suppl II):II-4-II-6
Measurement of motor recovery after stroke. Outcome assessment and sample size requirements.
P W Duncan, L B Goldstein, D Matchar, G W Divine and J Feussner

Stroke. 1992;23:1084-1089
doi: 10.1161/01.STR.23.8.1084

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/23/8/1084