A New Method for Predicting Recovery After Stroke

Kate Tilling, PhD; Jonathan A.C. Sterne, PhD; Anthony G. Rudd, FRCP; Thomas A. Glass, PhD; Robert J. Wityk, MD; Charles D.A. Wolfe, MD

Background and Purpose—Several prognostic factors have been identified for outcome after stroke. However, there is a need for empirically derived models that can predict outcome and assist in medical management during rehabilitation. To be useful, these models should take into account early changes in recovery and individual patient characteristics. We present such a model and demonstrate its clinical utility.

Methods—Data on functional recovery (Barthel Index) at 0, 2, 4, 6, and 12 months after stroke were collected prospectively for 299 stroke patients at 2 London hospitals. Multilevel models were used to model recovery trajectories, allowing for day-to-day and between-patient variation. The predictive performance of the model was validated with an independent cohort of 710 stroke patients.

Results—Urinary incontinence, sex, prestroke disability, and dysarthria affected the level of outcome after stroke; age, dysphasia, and limb deficit also affected the rate of recovery. Applying this to the validation cohort, the average difference between predicted and observed Barthel Index was −0.4, with 90% limits of agreement from −7 to 6. Predicted Barthel Index lay within 3 points of the observed Barthel Index on 49% of occasions and improved to 69% when patients’ recovery histories were taken into account.

Conclusions—The model predicts recovery at various stages of rehabilitation in ways that could improve clinical decision making. Predictions can be altered in light of observed recovery. This model is a potentially useful tool for comparing individual patients with average recovery trajectories. Patients at elevated risk could be identified and interventions initiated. (Stroke. 2001;32:2867-2873.)

Key Words: activities of daily living ■ disability evaluation ■ models, statistical ■ prognosis ■ recovery of function

A primary concern immediately after stroke for patients, their relatives, and their caregivers is the prospect for recovery. To address this concern and to aid in clinical management, individual predictions of recovery are needed; information about the “average” recovery pattern may have little relevance to an individual clinician or the patient. Prognostic studies of outcome after stroke have tended to concentrate on predicting outcome at a specific time point such as 6 months after stroke.1 This type of prediction does not aid clinical decisions about whether to continue an intervention such as a rehabilitation program or the identification of causes of failure to recover. The patient’s rate of recovery also has important implications for costs of care, especially the length of hospital stay.2

Many studies of recovery after stroke3–5 have estimated the probability of a patient achieving a specific benchmark—eg, independence in a number of activities of daily living—by a given time point—eg, discharge from stroke rehabilitation.6 Such probabilities may vary according to treatment regimen or patient characteristics. This work has not yielded clinically useful algorithms because this approach does not allow for individually tailored prediction of recovery based on particular characteristics of the patient and important early markers of progress. In addition, no information is available on the actual level of recovery if the benchmark is not achieved.

It has been emphasized that normal patterns of recovery from stroke should be established as a guide to monitoring recovery of future patients.3,7 However, carefully designed multivariable statistical models needed for the development of these normal patterns of recovery have been lacking. In a previous article, we presented the statistical theory and justification for a new approach to modeling recovery over time.8 Here, we demonstrate its clinical utility.

Longitudinal outcome data require nonstandard statistical analyses because repeated measures on the same individual are not independent, which is an assumption of standard statistical models.9 The number of patients in the study may decrease over time, and any model needs to be based on all
patients, not just those who survive until the end of the study. These issues can be addressed with multilevel models, which take into account the fact that the data are nested as observations (first level) within individuals (second level). The aim of this study was to demonstrate the clinical utility of our newly developed model that takes these statistical issues into account. This model can be used to make clinically useful predictions of the course of functional recovery based on patient characteristics and initial progress. We used data from 2 sources—a clinical trial of a rehabilitation program and a second, independent observational cohort—of stroke patients to examine the accuracy of these predictions.

Subjects and Methods

The model development cohort comprised 299 stroke patients recruited between January 1993 and July 1995 to a randomized trial (described elsewhere) and followed up at least once. This trial compared conventional management and supported early discharge from hospital with intensive rehabilitation therapy. It was conducted in accordance with institutional guidelines and approved by an ethics committee. To enter the trial, patients had to be able to transfer from bed to chair independently or with the help of their usual caregiver. Therefore, this cohort had relatively good physical mobility compared with an unselected group of stroke patients. We also used data from a second, independent cohort of patients to examine the accuracy and precision of predictions (see Validation of Predictions).

Patient characteristics, including age, sex, ethnicity (white/not white), and prestroke disability (defined as a Rankin Scale score of ≤3), were collected within 48 hours of randomization. Data on the severity of stroke had been recorded within 2 weeks of the stroke and were extracted from case notes. These included the presence of urinary incontinence, dysarthria, limb deficit (weakness or paralysis in any limb), swallowing deficit, dysphasia, and level of consciousness (with the Glasgow Coma Scale), all measured at the time of maximum impairment. All stroke patients, including both ischemic and hemorrhagic stroke patients, were eligible for inclusion.

Functional recovery was assessed at randomization and at ∼2, 4, 6, and 12 months after randomization with the Barthel Index (BI). The scoring of the BI is somewhat different in the United Kingdom than in the United States. Dependence in a given activity scores 0, with degrees of independence scoring 1, 2, or 3, depending on the activity (corresponding to scores of 5, 10, or 15 in the US BI). The total score has a minimum of 0 (completely dependent) and a maximum of 20 (completely independent).

As previously reported, there was no evidence for an effect on the main outcome for this trial (BI at 12 months after randomization). The mean age for this cohort was 70 years (SD=11.1 years); 165 (53%) were male; and 219 (73%) were white. The baseline characteristics of this cohort are shown in Table 1.

Statistical Methods

Multilevel growth curve models were used to estimate the effect on the BI of patient characteristics and stroke severity variables at baseline and how these effects varied over time. A multilevel growth curve model predicts the average recovery curve for patients with a given set of characteristics. It is assumed that each patient has his or her own true recovery curve and that these true recovery curves vary about the average recovery curve. Thus, the models allow recovery patterns to differ, even among patients with the same baseline characteristics. In addition, the observed outcomes for each patient will vary about their true recovery curve because of measurement error and day-to-day variation. In addition to estimating the average recovery curve and how this curve is related to patient characteristics, the multilevel growth curve also quantifies the variation between individual patient recovery curves and the degree of variability because of measurement error and day-to-day variation. Technical aspects of the models are given in detail elsewhere.

The first model fitted included all the patient characteristics listed in the previous section, with subtype of stroke defined as ischemic, hemorrhagic, or ill defined/unclassified. Variables were selected for inclusion in the final model by use of backward selection; a variable was retained in the model if the probability value for its effect (using a likelihood ratio test) was <0.1. The same criterion was used to assess whether the effect of each variable changed with time since stroke. The model describes the relationship between BI and time since stroke and thus can be used to predict recovery at any time point up to 1 year after stroke.

The parameters of the model can be used to estimate the correlation between the outcomes for a particular patient at any 2 time points. The reason is that the model accounts for both measurement error and day-to-day variability and the degree to which patients’ recovery curves differ. Given the observed outcome for an individual at 1 time point, this correlation can be used to predict the outcome for that individual at a future time point. Thus, predictions of an individual’s future pattern of recovery can be made after their characteristics at the time of stroke and their recovery history to the present are taken into account. This allows predictions for any individual to be modified in light of their observed recovery without any change to the underlying model.

Multilevel models were fitted by use of MLn, and all other statistical analyses were performed with Stata.

Validation of Predictions

The accuracy and precision of the predictions from the model were evaluated by use of data from an observational cohort of 1140 stroke patients registered by the South London Stroke Register between January 1995 and December 1998. This cohort of patients had the same baseline covariates measured as the trial cohort. The South London Stroke Register aims to register all first strokes in a geographically defined population; thus, this unselected cohort is representative of the entire stroke population in this area. Comparison of this cohort and the Framingham cohort of 501 stroke patients showed them to have comparable mean age at stroke (74 years in the South London Stroke Register, 72 years in the Framingham cohort), proportion of men (48% versus 43%, respectively), and mean acute BI (9.6 versus 10.6).

Only those 710 patients with measurements of BI on ≥2 occasions after stroke were included in the validation exercise. The mean age of this cohort was 70 years (SD=13.5 years); 374 (53%) were male; and 546 (77%) were white. The baseline characteristics of this cohort are shown in Table 1. The model was used to produce predictions of BI, the first based only on patient characteristics at the time of stroke.
and the second additionally on patients’ BI scores measured at the previous visit. Limits of agreement between predicted and actual BI scores were calculated as a way to validate the clinical utility of this model.18

Results

Pattern of Recovery After Stroke

The 299 subjects, 30 of whom died before the end of the study, were assessed on a total of 1346 occasions. Table 2 shows the mean BI at each occasion for all patients and for those who were and were not assessed at every occasion. Mean BI increased initially and then decreased slightly between 4 months and 1 year. The average pattern of recovery was an initial increase in BI, then a longer-term plateau, and then a decline (Figure 1).

Factors Associated With Recovery After Stroke

Presence of prestroke disability, urinary incontinence, dysarthria, and sex were associated with lower BI after stroke (Table 3), but their effects did not vary over time. For example, a stroke patient with dysarthria at baseline would be expected to have a BI =1 point lower than a stroke patient with no dysarthria at baseline at any time after stroke. The likely size of some of these differences could be modest at the individual level (eg, the effect of dysarthria could be as small as a mean decrease of 0.38 points on the BI) but potentially important at the population level.

There was no evidence of an effect of intervention (ie, intensity of rehabilitation therapy), ethnic group, or ischemic versus hemorrhagic stroke. The lack of a significant difference in outcome between hemorrhagic and ischemic strokes may be due to low power because there were few hemorrhagic strokes (9%).

The effects of age, dysphasia, and limb deficit at baseline on BI varied over time (Figure 2). On average, patients >80 years of age, with dysphasia or a limb deficit, had poorer functional recovery after stroke. In addition, patients >80 years of age tended to improve faster than younger patients initially but then to show a sharper long-term decline. Patients with dysphasia tended to improve faster initially, but after ~12 weeks after stroke, the recovery curves for those with and without dysphasia appeared to be parallel. Those with no limb paralysis or weakness showed little recovery because of the “ceiling” effect of the BI (Figure 2). Those with limb deficit improved quickly initially but then showed a slightly steeper long-term decline than those with no limb deficit.

There were also clear differences in patterns of BI over time between patients who died during the study and those who survived. Patients who died showed a considerably more rapid decline in outcome than those who survived (Figure 2).

Prediction of Outcome Based on Baseline Patient Characteristics

Patient characteristics (age, sex, prestroke disability), together with the stroke severity measurements (taken at time of maximum clinical impairment), can be used to predict BI at any time up to 1 year after stroke. For example, the predicted BI at 8 weeks after stroke for a man <80 years of age with limb deficit, urinary incontinence, and dysarthria is...
15.5. This information could be used to plan initial rehabilitation targets, make preliminary predictions about length of stay in hospital, and provide patients and their caregivers with some information on the likely pattern of recovery.

**Prediction of Outcome Based Additionally on Recovery History**

The model parameters can be used to derive correlation coefficients for the prediction of outcome at 1 time point with outcome measured at a previous time point. These can be used to alter the predictions for future outcome by taking into account the current observation of outcome. For example, suppose that a man <80 years of age with limb deficit, urinary incontinence, and dysarthria had a BI of 6 when he was assessed at 8 weeks after stroke (much lower than the predicted value of 15.5). This is taken into account in predictions of his outcome at 20 weeks after stroke, giving a predicted value of 8.9 instead of the prediction based on patient characteristics alone, which is 16.

Information such as this could be clinically useful to determine the likelihood of success in rehabilitation. For example, many patients are discharged from acute rehabilitation if they fail to meet rehabilitation goals. This model could be used to identify patients who are likely to have better outcome even if they demonstrate an early recovery plateau. These patients might otherwise not be eligible for continued access to rehabilitation services. On the other hand, the model could be used to identify patients who are unlikely to make good progress despite an absence of clear indicators of poor prognosis.

Details on the use of this model to predict values of the BI for individual patients are available from the authors. These calculations can also be made directly with a simple calculation that uses patient characteristics and the estimated model parameters. An Excel spreadsheet that does this is available from the authors at [http://www.epi.bris.ac.uk/user/jons/stroke](http://www.epi.bris.ac.uk/user/jons/stroke).

**Validation of Prediction of Recovery After Stroke**

We used an independent, unselected cohort of patients from a stroke registry in London to investigate the ability of the model to predict individual patients’ recovery patterns. BI was measured on a total of 1959 occasions for 710 patients: at 7 days (all patients), discharge from hospital (110 patients), and 3 months (692 patients) and 1 year (447 patients) after stroke. Sixty-five patients died before the end of the study. The mean BI over these 1959 occasions was 14.1 (SD=6.44). For predictions based only
on patients’ baseline characteristics and their stroke severity (1959 separate patient occasions at the time points described above), the average predicted BI was 15.4 (SD=2.57), with a minimum predicted value of 6 and a maximum of 20. The average difference between the predicted and actual BI scores was 1.27 (SD=5.43), with 90% limits of agreement from −8 to 10. On 936 occasions (49%), the predicted BI was within 3 points of the measured BI.

Prediction was greatly improved when the measurement for a subject was also used to predict BI at the next occasion. There were 1242 measurements of BI with an immediately previous measurement (eg, a 1-year BI measurement with a 3-month BI measurement for the same individual) at hospital discharge and 3 months and 1 year after stroke. The average BI for these 1242 occasions was 15.5 (SD=5.62). The predicted BI was on average 15.1 (SD=4.43), with a minimum of 0 and a maximum of 20. For these 1242 occasions, the average difference between the predicted and actual BI scores was −0.4 (SD=3.84), with 90% limits of agreement from −7 to 6. Conditionally predicted BI was within 3 points of measured BI on 859 occasions (69%).

Prediction of Death After Stroke
Of the patients in the validation data set, 65 did not survive until 1 year after stroke. Because death tended to be preceded by a rapid decline in BI, a measured BI of >1 point below the predicted value was used to predict death before the next measurement occasion. Application to the validation data set (710 patients, 1242 measurements) yielded a sensitivity of 65% and specificity of 79% (42 of 65 measurements when the patient died before the next visit and 92 of 1177 when the patient survived).

Discussion
We developed a new prognostic model for functional recovery after stroke as measured by the BI. This model can be used to predict recovery at any time point after stroke. This approach is superior to previous models that can predict recovery at 1 time point only. The accuracy of predictions was increased when each patient’s observed recovery was taken into account when predictions were made. This model allows evolving, individualized predictions of recovery to be made for any stroke patient. Such predictions can assist clinicians and families in the course of poststroke rehabilitation.

The predictions made on the basis of each individual’s observed recovery history for the validation data set were of sufficient accuracy (69% of predictions lay within 3 points of the true value) to be of use in a clinical setting. This is especially important given that, at present, there is no way to make predictions of outcome over time since stroke or to include knowledge of previous outcome for a patient in predictions of future outcome. Rehabilitation health professionals could use this model to monitor the recovery of stroke patients. Initial predictions of recovery could be used to set rehabilitation targets. Actual recovery for each patient could be compared with that predicted by the model, and this information could be used to adjust the amount of rehabilitation therapy received. The comparison of actual and predicted BI scores could also be used to identify patients who were not recovering as well as expected. For these patients, additional intervention may be warranted.

This study has 2 important weaknesses. First, no comprehensive neurological scale such as the Scandinavian Stroke Scale or the NIH Stroke Scale was available for this cohort. However, many of the variables included in these scales, eg, level of consciousness, speech, and swallowing deficits, were also included here as covariates. The major prognostic factors for which data were unavailable in this study were orientation, gaze and visual field, sensory ability, and neglect. Despite omitting these factors, the model still showed reasonable predictive accuracy on an independent cohort of patients.

Second, the model was developed initially with a selected cohort of stroke patients drawn from a clinical trial of a rehabilitation program who met minimum mobility and self-care criteria. This was reflected in the predicted values for the independent cohort, which were too high (ie, predicted disability was too low) when based only on patient characteristics at the time of stroke. However, there was no evidence of bias in the predictions based additionally on recovery history, indicating that the selection of the development cohort affected mean outcome but not the pattern of recovery or the relationship between outcome at different time points. In addition, only the extremely severe strokes will not be represented in the model development (trial) cohort, and these stroke patients may be very different from most stroke patients. Any selection biases associated with the use of this cohort are likely to be less important given that more severely impaired patients do not normally make use of rehabilitation and thus the prognostic model may be of little relevance to them.

The model was developed with a selection of baseline characteristics, including stroke severity indicators. Many other factors may influence recovery, including stroke subtype, intensity and type of therapy, and psychosocial factors. However, the aim of this model was to use simple, easily measured variables to predict recovery to make the model as generalizable as possible. These methods could be extended to include variables of interest to specific groups by developing more complex models.

Although many studies have evaluated outcome after stroke, to the best of our knowledge, none has quantified patterns of recovery over time in this way. The general pattern of initial rapid recovery followed by a plateau that we have shown corroborates several other descriptive studies. A trial of physiotherapy late after stroke also demonstrated a long-term decline in outcome after stroke, which could be reversed by physiotherapy intervention. We have extended previous work in this area by demonstrating the predictive value of a statistical model that takes into account individual characteristics and previous progress. This model allows the average functional recovery over the entire period up to 1 year after stroke to be quantified. Beyond this point, the BI should be supplemented with more appropriate measures of long-term recovery, which may be more sensitive to small improvements and to adaptations made by the patient to overcome any residual disabilities or impairments.
We have separated patient characteristics at the time of stroke into those affecting average level of outcome (sex, prestroke disability, dysarthria, urinary incontinence) and those related to recovery rate (age, dysphasia, limb deficit). We found no evidence of any association between stroke subtype and outcome, possibly because there were few hemorrhagic strokes (9%). Previous studies have shown that urinary incontinence, increasing age, prestroke disability, and presence of motor impairment (all measured at the time of maximum impairment) are associated with a poorer outcome after stroke. However, estimation of the effect of characteristics on rate of recovery was not possible with simpler, cross-sectional methods.

As with any standard linear regression model, our model ignores the discrete nature of the BI. However, predicted scores can be rounded to the nearest whole number, and in our example, this made negligible difference to the predictive power and validity. In addition, the BI displays both “floor” and “ceiling” effects. For example, recovery may continue once a person has achieved a BI score of 20, but the scale will be unable to capture it (ceiling effect). This problem could be addressed by supplementing the BI with a measure of extended activities of daily living, eg, the Frenchay Activities Index.

Our results have implications for the design and analysis of stroke intervention trials. Figure 3 shows a graphical representation of the effect of intervention on recovery. The end point of interest, the difference in outcome between the two arms of the trial, differs according to the time since stroke. For example, if outcome were compared only at 6 months, we would be unable to capture it (ceiling effect). This problem could be addressed by supplementing the BI with a measure of extended activities of daily living, eg, the Frenchay Activities Index.

We conclude that intervention had little or no effect, although in this manuscript.

References


A New Method for Predicting Recovery After Stroke
Kate Tilling, Jonathan A.C. Sterne, Anthony G. Rudd, Thomas A. Glass, Robert J. Wityk and Charles D.A. Wolfe

Stroke. 2001;32:2867-2873
doi: 10.1161/hs1201.099413
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2001 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/32/12/2867