Predicting Outcome After Acute and Subacute Stroke
Development and Validation of New Prognostic Models

Carl Counsell, MD; Martin Dennis, FRCP; Michael McDowall, MSc; Charles Warlow, FRCP

Background and Purpose—Statistical models to predict the outcome of patients with acute and subacute stroke could have several uses, but no adequate models exist. We therefore developed and validated new models.

Methods—Regression models to predict survival to 30 days after stroke and survival in a nondisabled state at 6 months were produced with the use of established guidelines on 530 patients from a stroke incidence study. Three models were produced for each outcome with progressively more detailed sets of predictor variables collected within 30 days of stroke onset. The models were externally validated and compared on 2 independent cohorts of stroke patients (538 and 1330 patients) by calculating the area under receiver operating characteristic curves (AUC) and by plotting calibration graphs.

Results—Models that included only 6 simple variables (age, living alone, independence in activities of daily living before the stroke, the verbal component of the Glasgow Coma Scale, arm power, ability to walk) generally performed as well as more complex models in both validation cohorts (AUC 0.84 to 0.88). They had good calibration but were overoptimistic in patients with the highest predicted probabilities of being independent. There were no differences in AUCs between patients seen within 48 hours of stroke onset and those seen later; between ischemic and hemorrhagic strokes; and between those with and without a previous stroke.

Conclusions—The simple models performed well enough to be used for epidemiological purposes such as stratification in trials or correction for case mix. However, clinicians should be cautious about using these models, especially in hyperacute stroke, to influence individual patient management until they have been further evaluated. Further research is required to test whether additional information from brain imaging improves predictive accuracy. (Stroke. 2002;33:1041-1047.)

Key Words: cerebrovascular disorders ■ models, statistical ■ prognosis

Accurate prognostic models for patients with acute and subacute stroke would have several important uses. They could guide patient management1-2 (eg, patients with a good prognosis could be spared potentially risky treatments such as thrombolysis); allow more reliable information to be given to patients and their relatives2; improve planning of rehabilitation and discharge for patients3; correct for case mix when different cohorts of patients are compared (eg, in hospital “league” tables)3-4; and be used to analyze the results of randomized trials or meta-analyses according to baseline severity.5-6 Unfortunately, none of the existing models have been adequately generated, and very few have been properly validated.7 The accuracy of a model’s predictions must be externally validated in at least one (preferably several) independent cohort of patients (test cohort) that was not used to generate the model.1-2,8 We developed new prognostic models according to established guidelines1-3,8,9 and then validated them in 2 independent cohorts of stroke patients.
Definition of Outcomes

We were interested in 2 important but simple outcomes: survival at 30 days and survival in an independent state (Oxford Handicap Scale score $\geq 3$) at 6 months.3,13,14 Models for outcomes at 1 year are available from the authors on request. Outcomes were available for all 530 patients and were assessed blind to the predictor variables.

Data Reduction and Definitions of Predictor Variables

Predictor variables must be easy to collect (to minimize missing data), clinically relevant, and reliable.1,3,8,9,15 The number of variables in multiple regression analyses must also be carefully controlled.9,16 Too few variables means that important predictors may be omitted, while too many variables can result in overfitting (a type I error in which false-positive predictors are erroneously included in the model); underfitting (a type II error in which important variables are omitted from the final model); and paradoxical fitting (a type III error in which a variable that, in truth, has a positive association with the outcome is found to have a negative association).16,17 The risk of these problems increases as the ratio of outcome events to the number of predictor variables becomes smaller (the events per variable [EPV] ratio, in which the number of events is the lower figure for binary outcomes). The risk of error is especially high with EPVs $\leq 10$.17

In the OCSP, 131 baseline variables were initially collected, whereas the optimal number to ensure an EPV of $\geq 35$ was approximately 25. We reduced the number of variables by excluding (1) those with too many missing data (ie, on $>50$ patients) because this suggested that they were not easy to collect; (2) those with too little data (predictive factor present in $<10$ patients); (3) those that we considered unlikely to be associated with the outcome (eg, whether the onset occurred during sleep); and (4) those with questionable reliability20 (eg, sensory assessments). Data from CT scans were excluded because 186 patients (35%) did not have a scan or were scanned $>2$ weeks after their stroke, and the early generation scans available provided limited data compared with modern CT. Several of the remaining variables were combined to make composite variables; this resulted in a final set of 39 variables (Appendix). This number was higher than the 25 required to maintain the EPV at

<table>
<thead>
<tr>
<th>Variable (% Not Assessable in OCSP*)</th>
<th>OCSP (n=530)</th>
<th>SEPIVAC/Perth (n=538)</th>
<th>LSR (n=1330)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alive at 30 d Model Set 1</td>
<td>Set 1 Model</td>
<td>Set 1 + 2 Model</td>
<td>Set 1 Model</td>
</tr>
<tr>
<td>Alive and Independent at 6 mo</td>
<td>Model Set 1</td>
<td>Model Set 1 + 2 + 3</td>
<td>Model Set 1</td>
</tr>
<tr>
<td>Age, mean [SD], y (0)</td>
<td>73 [12]</td>
<td>73 [12]</td>
<td>68 [13]</td>
</tr>
<tr>
<td>Male (0)</td>
<td>48</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>Living alone (1)</td>
<td>46</td>
<td>36</td>
<td>31</td>
</tr>
<tr>
<td>Employed at time of stroke (0)</td>
<td>16</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>Independent before stroke (0)</td>
<td>85</td>
<td>85</td>
<td>87</td>
</tr>
<tr>
<td>Previous myocardial infarction (0)</td>
<td>17</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Known diabetic before stroke (0)</td>
<td>8</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>Previous malignancy (0)</td>
<td>10</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Normal Glasgow Coma Scale eye score (0)</td>
<td>90</td>
<td>86</td>
<td>91</td>
</tr>
<tr>
<td>Normal Glasgow Coma Scale motor score (0)</td>
<td>91</td>
<td>88</td>
<td>94</td>
</tr>
<tr>
<td>Normal Glasgow Coma Scale verbal score (0)</td>
<td>79</td>
<td>64</td>
<td>77</td>
</tr>
<tr>
<td>Able to lift both arms (0)</td>
<td>76</td>
<td>70</td>
<td>78</td>
</tr>
<tr>
<td>Able to walk (0)</td>
<td>68</td>
<td>50</td>
<td>51</td>
</tr>
<tr>
<td>Current smoker (3)</td>
<td>26</td>
<td>20</td>
<td>36</td>
</tr>
<tr>
<td>Cardiac disease (0)</td>
<td>29</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>Cognitive deficit (20)</td>
<td>35</td>
<td>49</td>
<td>33</td>
</tr>
<tr>
<td>Visual field defect (11)</td>
<td>30</td>
<td>36</td>
<td>34</td>
</tr>
<tr>
<td>Normal proprioception (12)</td>
<td>64</td>
<td>72</td>
<td>69</td>
</tr>
<tr>
<td>Anemia (1)</td>
<td>7</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Urea $\leq$ 7 mmol/L (2)</td>
<td>59</td>
<td>22</td>
<td>64</td>
</tr>
<tr>
<td>Glucose $&lt;7$ mmol/L (3)</td>
<td>74</td>
<td>86</td>
<td>76</td>
</tr>
<tr>
<td>Glucose $&gt;11$ mmol/L (3)</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Alive at 30 days</td>
<td>87</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Alive and Independent at 6 months</td>
<td>51</td>
<td>33</td>
</tr>
</tbody>
</table>

+ indicates independent predictor of outcome.

*Percentage of patients in training cohort in whom data was missing or unassessable e.g. because they were drowsy/confused/dysphasic.
†Except for age.
Validation of Models

We had access to 3 independent prospective cohorts (test cohorts) of stroke patients that had been established to collect data similar to those of the OCSP. Two of these cohorts were community based. SEPIVAC (Italian initials for Epidemiological Study of Incidence of Acute Cerebrovascular Disease) collected data on first-ever strokes from a region in Italy (1986–1989), while the Perth Community Stroke Study collected data on all strokes in Perth, Australia (1989–1990). The median time from stroke onset to assessment in both cohorts was 4 to 5 days. These cohorts had relatively small numbers of eligible patients (310 in SEPIVAC, 228 in Perth), and therefore we combined them to make 1 community-based cohort that was restricted to patients with a first-ever stroke (not SAH) who were assessed by a study neurologist within 30 days of onset. The other test cohort, the Lothian Stroke Register (LSR), was hospital based and included stroke patients (both inpatients and outpatients, first and recurrent strokes excluding SAH) from 1 hospital in Edinburgh who were prospectively registered, examined, and followed up from 1990 onward. For validation, we again restricted this cohort to those seen within 30 days of onset.

We plotted calibration curves of the proportion of patients in each test cohort who actually had a good outcome against the proportion predicted by the model (in deciles of predicted probability). We assessed discrimination by calculating the area under a receiver operating characteristic (ROC) curve (AUC) of sensitivity versus 1 minus specificity. An area of 1 implies a test with perfect sensitivity and specificity, while an area of 0.5 implies that the model’s predictions are no better than chance. The AUC for each model was calculated with the trapezoidal rule.

Because each outcome had 3 models (1 developed from set 1 variables, 1 from set 1 and set 2 variables, and 1 from all variables), the best model for each outcome in each test cohort was defined as the model with the largest AUC or, if there were no statistically significant differences in areas, the model with the simplest variables. Since we made multiple comparisons between different ROC curves, we chose a more rigorous definition of statistical significance (2P≤0.01). For the logistic regression model for alive and independent, we also produced variance/covariance matrices so that confidence intervals could be produced for each prediction.

Using the best model (as defined above) in the hospital test cohort (LSR), we determined whether the AUC (ie, accuracy of the predictions) for alive and independent at 6 months varied according to the following: whether the patients were inpatients or outpatients; and whether or not the patients had had a previous hemorrhagic stroke verified by CT scanning within 14 days of onset versus the rest); whether the patients were seen as inpatients or outpatients; and whether or not the patients had had a previous stroke. These subgroup analyses were restricted to the LSR test cohort because this cohort had the most data.

Results

New Models

The characteristics and outcomes of the patients in the OCSP are given in Table 1 (only variables that appeared in any of the models are shown), along with the percentage of patients with data that were not assessable in the training cohort and those variables that were independent predictors in each model. Analyses showed that the models did not violate any of the necessary statistical assumptions. The simplest model for alive and independent at 6 months only included 6 variables (age, living alone, prior disability, the verbal component of the
Glasgow Coma Scale, arm power, and ability to walk), and these same variables also appeared in most other models. We therefore generated 1 additional survival model using only these 6 variables and compared this with the original set 1 survival model that included 11 variables.

Validation
The characteristics of patients in the 2 test cohorts are shown in Table 1. These were broadly similar to the OCSP cohort, although the OCSP included patients with slightly less severe strokes than the community test cohort, and patients in the hospital test cohort were 5 years younger on average. Fewer patients from the community test cohort were alive and independent at 6 months compared with the OCSP.

The AUCs for each model in each test cohort are shown in Table 2, and the ROC curves for the hospital test cohort are shown in Figure 1. In both test cohorts, the models using only set 1 variables generally did not have significantly smaller AUCs than the more complex models (indeed, they often gave slightly larger areas). The only exception was alive and independent in the hospital test cohort, where the set 1 model was marginally better ($P=0.01$) than the set 1 model. Similarly, the survival model with only 6 simple variables was as good as the set 1 model with 11 variables. The exact models using only the 6 simple variables are given in Table 3.

The predictions of the simplest model for each outcome were reasonably well calibrated against the actual outcomes (Figure 2). The model for alive and independent overestimated the proportion who were independent in those with the highest predicted probabilities ($>0.7$) of good outcome. The reason for this was unclear.

The comparisons of the predictions of the best model (set 1 + 2) for alive and independent at 6 months in various different types of stroke patient in the LSR are shown in Table 4. There were no clear differences in the predictions between the various subgroups except that the discrimination was poorer for patients first seen as outpatients than for inpatients. The actual prognosis in outpatients was, as expected, much better than in the inpatients: $>99\%$ of outpatients were alive at 30 days, and $82\%$ were alive and independent compared with $86\%$ and $37\%$, respectively, in the inpatients. Outpatients were also seen much later after their stroke onset (median time, 11 days) than inpatients (median time, 2 days). Similar results for all these subgroup analyses were found with the simple set 1 models and if survival alone was analyzed.

**Discussion**
These new models have several strengths. They were developed from high-quality data following established guidelines and have good predictive accuracy in 2 independent cohorts from different settings, countries, and time periods. The accuracy in the LSR is particularly reassuring because this was a much more recent cohort and therefore patients in this cohort were often treated with early administration of aspirin in a stroke unit. Because the models were developed on a community-based sample of patients with any type of stroke, they are more likely to be generalizable to other cohorts. We have demonstrated validity in those with both ischemic and hemorrhagic strokes (although our definition of hemorrhagic may have missed some small hemorrhages in patients not scanned within the first few days), in those with first-ever and recurrent stroke, and in those seen within 48 hours of stroke onset and those seen later. However, although there was a nonsignificant trend for better predictions in those seen within 48 hours, very few patients were seen within 6 hours of onset, and therefore the accuracy of the models in those seen in the hyperacute phase of stroke is unknown. The models may also be less applicable to very young patients (only 5% to 10% of patients in the training and test cohorts were aged $<50$ years), and they have poor discrimination in outpatients. The LSR outpatients were highly selected patients with a very good outcome, and the models were not able to predict the small numbers of outpatients who did poorly.

There were several other limitations of this study. The numbers of patients and outcomes on which the models were generated, although much larger than most comparable studies,$^7$ were still relatively small, and many different regression analy-
TABLE 3. Simplest Models for Predicting Survival and Alive/Independent

<table>
<thead>
<tr>
<th>Variable*</th>
<th>Definition of Variable</th>
<th>Parameter, b (SE)</th>
<th>Relative Risk of Survival (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age at onset in years to 2 decimal places</td>
<td>0.034 (0.007)</td>
<td>0.975 (0.95–0.98)</td>
</tr>
<tr>
<td>Living alone</td>
<td>Nobody permanently living with patient before stroke</td>
<td>−0.406 (0.140)</td>
<td>0.67 (0.51–0.88)</td>
</tr>
<tr>
<td>Independent before stroke</td>
<td>Oxford Handicap Scale score ≥2 before stroke</td>
<td>0.501 (0.166)</td>
<td>1.65 (1.19–2.29)</td>
</tr>
<tr>
<td>Normal Glasgow Coma Scale verbal score</td>
<td>Glasgow Coma Scale verbal score=5 (ie, oriented)</td>
<td>0.776 (0.153)</td>
<td>2.17 (1.61–2.93)</td>
</tr>
<tr>
<td>Able to lift arms</td>
<td>Able to lift both arms to horizontal (MRC score ≥3 both arms)</td>
<td>0.851 (0.171)</td>
<td>2.34 (1.68–3.27)</td>
</tr>
<tr>
<td>Able to walk</td>
<td>Walks without the help of another person (can use stick/frame)</td>
<td>0.489 (0.178)</td>
<td>1.63 (1.15–2.31)</td>
</tr>
<tr>
<td>Baseline survival (S_0(t)) at 30 days = 0.631</td>
<td>(\text{Parameter, } b \ (SE))</td>
<td>(\text{Odds Ratio (95% CI)})</td>
<td></td>
</tr>
<tr>
<td>Logistic regression model for alive and independent at 6 mo$‡$</td>
<td>(\text{Coefficient, } b \ (SE))</td>
<td>(\text{Coefficient, } b \ (SE))</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>12.340 (1.482)</td>
<td>3.71 (1.77–7.77)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>−0.051 (0.013)</td>
<td>0.955 (0.93–0.97)</td>
<td></td>
</tr>
<tr>
<td>Living alone</td>
<td>0.661 (0.261)</td>
<td>0.52 (0.31–0.86)</td>
<td></td>
</tr>
<tr>
<td>Independent before stroke</td>
<td>−2.744 (0.514)</td>
<td>15.55 (5.68–42.58)</td>
<td></td>
</tr>
<tr>
<td>Normal Glasgow Coma Scale verbal score</td>
<td>−2.160 (0.473)</td>
<td>8.67 (3.43–21.91)</td>
<td></td>
</tr>
<tr>
<td>Able to lift arms</td>
<td>−2.106 (0.473)</td>
<td>8.22 (3.25–20.76)</td>
<td></td>
</tr>
<tr>
<td>Able to walk</td>
<td>−1.311 (0.377)</td>
<td>3.71 (1.77–7.77)</td>
<td></td>
</tr>
</tbody>
</table>

MRC indicates Medical Research Council.
*Dichotomous variables were coded as follows: 1 = Yes, 2 = No.
†Probability of survival \(S(t)\) is calculated by \(S(t) = S_0(t)^{eY}/(1 + eY)\), where \(S_0(t)\) is baseline survival at time \(t\) and \(Y = b_1X_1 + b_2X_2 + \ldots + b_nX_n\). \(X_i\) being the value of the \(i\)th variable and \(C\) the value chosen for calculation of the baseline survival. Baseline survival at 30 days was calculated for person aged 70 scoring 0.65, indicating excellent performance, other variables.
‡Probability of outcome is calculated by \(P = e^{\beta}/(1 + e^{\beta})\), where \(Y = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n\).
§Per year of age.

ses were performed. This may have resulted in false-positive variables in some models and in important variables being excluded from others because of lack of power.31 One hundred twelve patients were excluded from the training data set, but these patients were equally divided between those with severe strokes (ie, who died early before they could be seen) and those with milder strokes (ie, survivors who were first seen >30 days after their stroke). Few of these patients were admitted to the hospital (they died too quickly or had very minor strokes), and therefore most would also have been excluded from hospital-based studies, except in countries where almost all subjects with suspected stroke or transient ischemic attack are admitted. Several potentially important variables were not included in the models. Urinary incontinence has been shown in several studies to be an important predictor of poor outcome,32,33 but this information was not collected for most patients in the OCSP because it predated these studies. Moreover, detailed brain imaging was not available for many patients in the OCSP, and therefore we were not able to analyze the predictive value of certain radiological variables on outcome (eg, size of infarct, presence of hemorrhage). However, this means that the models can be used in any healthcare system regardless of whether or not there is access to CT or MRI scanning.

We have demonstrated that the simplest models that included only 6 clinical variables usually gave predictions that were as good as (if not better than) models using more complicated, less easily collected variables. Although we had expected that including more detailed variables would improve the models’ performance, other researchers have also found that simple models predict as accurately as more complex ones.34–36 This may be because more complex variables are more difficult to measure reliably or because increasing the number of variables leads to overfitting. The 6 simple variables do have some face validity since they cover age, social circumstances, prior disability, and stroke severity in terms of both upper and lower limb function, reduced level of consciousness, cognition, and speech. An abnormal verbal component of the Glasgow Coma Scale may also indicate confusion and dysphasia as well as impaired consciousness, which may explain why this component was selected rather than the motor or eye components of the score. We have shown that the 6 variables can be collected reliably, with interobserver \(\kappa\) values all >0.65, indicating excellent agreement.37

The simplicity of these models will make them useful for epidemiological studies. We have already used them to correct for differences in case mix when outcomes of stroke patients from different hospitals were compared,38 to assess whether certain other variables were additional independent predictors of prognosis,39 and to stratify patients by baseline prognosis in trials.40,41

The use of models to influence the management of individual patients is more complex. They are certainly simple to use, either by
programming them into hand-held computers or using nomograms. However, they do not predict specific outcomes that may be more relevant in clinical management (eg, to determine whether the patient will regain useful speech or be able to walk). The models’ predictions would need to be better than physicians’ informal predictions and to improve patient care and outcomes. False-positive and false-negative predictions of good outcome could harm patients, eg, a patient falsely predicted to do poorly might be given a hazardous treatment or denied an effective treatment. Hence, we cannot recommend the models’ use in clinical practice until they have been evaluated in clinical trials.15

**Conclusions**

We believe that we have produced useful, straightforward, and valid models to predict outcome in most acute stroke patients in any inpatient clinical setting. Because the models depend on only 6 easily assessed and robust clinical variables, it takes just minutes to categorize individuals or groups of patients with differing prognosis. The models will be available interactively on www.dcn.ed.ac.uk/models. In the future, we plan to try to validate these models in other cohorts (eg, those seen within hours of stroke onset), to compare them with clinical predictions and other models, and to assess whether they can be improved by adding other variables (eg, urinary continence or imaging data).

**Appendix**

**Variables Entered Into the Multiple Regression Models**

**Set 1: Simple Clinical Variables**

Age, sex, living alone, employed, independent before stroke, history of hypertension, myocardial infarct, diabetes, malignancy, examined
within 2 days, systolic blood pressure >160 mm Hg, systolic blood pressure <120 mm Hg, Glasgow Coma Scale eye/motor/verbal scores, able to lift arms, able to lift legs, able to walk.

Set 2: More Detailed Clinical Variables
Current smoker, previous transient ischemic attack, peripheral vascular disease, apoplectic onset, cervical bruit, cardiac disease, dysphasia, cognitive deficit (eg, neglect, dyspraxia, or visuospatial problems), visual field defect, gaze palsy, brain stem function, proprioception.

Set 3: Investigation Results (Initial Results After Stroke)
High hemoglobin level (men >17 g/dL, women >16 g/dL), anemia (men <13 g/dL, women <11 g/dL), platelet count >400×10^9/L, platelet count <150×10^9/L, urea <7 mmol/L, glucose >11 mmol/L, glucose <7 mmol/L, any atrial fibrillation, abnormal cardiac rhythm.

Examination findings were obtained at initial assessment.

Acknowledgments
This study was supported by a Wellcome Trust training fellowship in clinical epidemiology (Dr C. Counsell) and by a Wellcome Trust project grant (M. McDowall). We thank the patients included in the various cohorts and the staff who collected the initial data; the principal investigators of the SEPIVAC and Perth studies who allowed us to use their data to validate the models (Dr Stephano Ricci for SEPIVAC, Dr Edward Stewart-Wynne, Dr Konrad Jamrozik, and Dr Craig Anderson for the Perth study); Robyn Broadhurst, who extracted the data from the Perth study to send to us; and Jim Slattery and Dave Signorini for statistical help and guidance.

References
Predicting Outcome After Acute and Subacute Stroke: Development and Validation of New Prognostic Models
Carl Counsell, Martin Dennis, Michael McDowall and Charles Warlow

Stroke. 2002;33:1041-1047
doi: 10.1161/hs0402.105909

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2002 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/33/4/1041

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/