

## Predicting the Long-Term Outcome after Subacute Stroke within the Middle Cerebral Artery Territory

Oh Young Bang, M.D., Ph.D., Hee Young Park, M.D., Jung Han Yoon, M.D.,  
Seung Hyeon Yeo, M.D., Ji Won Kim, Mi Ae Lee, Mi Hee Park,  
Phil Hyu Lee, M.D., Ph.D., In Soo Joo, M.D., Kyoong Huh, M.D.

*Department of Neurology, Ajou University School of Medicine, Suwon, Korea*

**Background and Purpose:** The National Institutes of Health Stroke Scale (NIHSS) score is known to be effective in predicting the likelihood of recovery after stroke. However, the baseline NIHSS score predicts long-term outcomes rather crudely because early changes in stroke scores may influence the stroke outcomes. Therefore, a precise prognostic algorithm or a cutoff point for predicting long-term outcomes based on data from serial NIHSS scores is needed.

**Methods:** We serially assessed 437 patients with acute symptomatic ischemic stroke within the middle cerebral artery territory who presented with nonlacunar stroke and were followed-up for at least 6 months after symptom onset. The NIHSS score was serially checked at 0, 1, 3, 7, and 14 days after admission. In all patients, the Barthel index (BI) and the modified Rankin Scale (mRS) score were checked, with a poor outcome defined as any of the following endpoints: death, modified mRS score of >3, or BI of <60.

**Results:** A marked neurological improvement or worsening (i.e., a change in the NIHSS score of at least 4) was seen in 13.5% or 5.5% of the patients, respectively, during the first 7 days after admission. About 25% of the 437 patients had poor long-term outcomes. Analysis of receiver operating characteristic curves showed that the NIHSS score at day 7 after admission was better for predicting poor long-term outcomes than was the baseline score ( $P=0.003$ ). In addition, we analyzed the cutoff point of the 7th-day NIHSS score for predicting a poor outcome at 6 months after symptom onset. An NIHSS score of at least 6 at day 7 after admission predicted poor long-term outcomes with a sensitivity of 84% [95% confidence interval (CI), 76-90%], a specificity of 92% (95% CI, 88-94%), and positive and negative predictive values of 77% and 95%, respectively. A logistic regression analysis revealed that age, diffusion-weighted imaging lesion volume, stroke history, and 7th-day NIHSS score were independently associated with poor outcome. However, no score used in addition to the 7th-day NIHSS score improved the prediction of a poor outcome.

**Conclusions:** An NIHSS score of at least 6 on day 7 after admission accurately forecasts a poor long-term outcome after stroke. Our data may be helpful in predicting the long-term prognosis as well as in making decisions regarding novel therapeutic applications in subacute-stroke trials.

*J Clin Neurol 1(2):148-158, 2005*

**Key Words :** Ischemic stroke, Outcome, Stroke scale

### INTRODUCTION

Forecasts of recovery from stroke have focused on the scores of neurological scales at admission or on the selection of patients to be treated with recombinant

Received : July 18, 2005 / Accepted : August 23, 2005 / Address for correspondence : Oh Young Bang, M.D., Ph.D.  
Department of Neurology, Ajou University School of Medicine, San 5 Woncheon-dong, Paldal-gu, Suwon, Gyeonggi-do, 442-749, Korea  
Tel: +82-31-219-5175, Fax: +82-31-219-5178, E-mail: nmboy@unitel.co.kr

tissue plasminogen activator.<sup>1-6</sup> Unfortunately, no models for subacute-stroke trials have been described. Accurate prognostic models for patients with subacute stroke would have several important uses, such as guiding patient management<sup>7,8</sup> (e.g., patients with a good prognosis could be spared potentially risky treatments such as stem cell therapy<sup>9,10</sup>), allowing more reliable information to be given to patients and their relatives,<sup>8</sup> and improving the planning of patient rehabilitation and discharge.<sup>11</sup>

The recovery profile of stroke can vary nonlinearly with the severity, with the course of improvement being greater at the lower end of the deficit scale. When this is the case, many patients enrolled with low predictive scores will improve regardless of treatment, reducing the likelihood of a real protective effect of treatment being detected. Conversely, some patients worsen relative to the initial severity of their stroke, resulting in further neuronal injury and less favorable outcomes.<sup>2</sup> False-positive and false-negative predictions of long-term outcome can be detrimental to patients, with the possibility of the needless provision of hazardous treatment or the incorrect denial of effective treatment, respectively. Therefore, understanding the clinical course after stroke and the time point for enrollment in subacute-stroke trials may help with the design of a prestudy randomization schedule and power analysis.

The National Institutes of Health Stroke Scale (NIHSS) is an attractive candidate predictor because it is widely used, is easily learned, and can be performed rapidly on admission. Moreover, the NIHSS score is known to predict the likelihood of patient recovery after stroke. However, the baseline NIHSS score predicts long-term outcomes rather crudely, and the fate of patients with intermediate scores is unclear.<sup>12</sup> Therefore, a precise prognostic algorithm or a cutoff point to predict poor long-term outcomes based on data from serial NIHSS scores is needed.

In this study, we evaluated the cutoff point - in terms of time point and severity as measured by serial NIHSS scores - for predicting long-term outcomes. In addition, we investigated a prognostic model for long-term outcomes.

## MATERIALS AND METHODS

### 1. Patient selection

From October 2000 to October 2004, we prospectively studied consecutive patients with acute symptomatic middle cerebral artery (MCA) territory infarcts who were admitted to the Department of Neurology at Ajou University Hospital, Korea. All patients had suffered from focal symptoms and had been observed within 72 hours of symptom onset, showed relevant lesions within the MCA distribution territory on magnetic resonance diffusion-weighted imaging (DWI), had undergone complete workups, including vascular and cardiologic workups, and were followed for more than 6 months (Fig. 1). All of the patients provided their informed consent to participate in the study. We excluded patients who presented with lacunar stroke or stroke recurrence within 6 months; those who were treated for thrombolysis, hypothermia, or craniectomy; and those who were discharged against medical advice, absconded from the hospital, or died during the acute phase of stroke, since such cases are not typical of the disposition after stroke care.

Of 748 stroke patients who were admitted during the study period, the above criteria lead to 437 patients being included in this study (Fig. 1): 263 men (60%) and 174 women (40%), aged  $62.1 \pm 12.8$  (mean  $\pm$

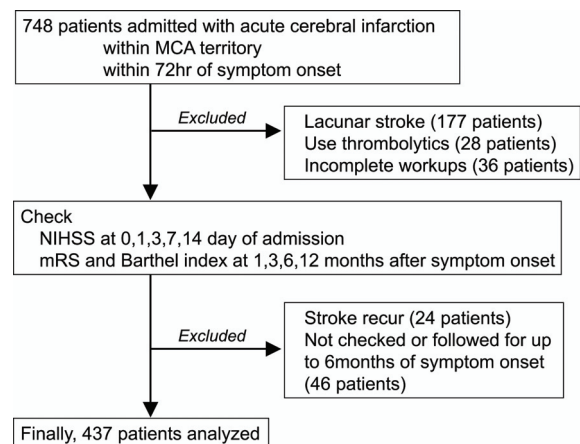


Figure 1. Patient selection.

SD) years (range, 32-91 years).

## 2. Workup

Patients were evaluated according to a protocol that included demographic data, medical history, vascular risk factors, and stroke scale scores [NIHSS score, Barthel index (BI), and the modified Rankin Scale (mRS) score], as in our previous study.<sup>13</sup> Both T2-weighted imaging and DWI were performed using a clinical 1.5-T magnetic resonance imaging system, and all patients underwent diagnostic testing, which included digital-subtraction or magnetic resonance angiography, electrocardiography, and routine blood tests. Echocardiography was performed in 381 patients (87%). The degree of stenosis was measured as in our previous study,<sup>13</sup> with occlusion defined as >50% stenosis or occlusion of the large intracranial vessels and internal carotid artery. We also measured the volumes of the DWI lesion(s) in 429 patients (98.2%). The DWI lesion volumes were calculated by multiplying the measured area per slice by the section thickness (conditions: repetition time, 10,000 ms; echo time, 104 ms; slice thickness, 7 mm; no gap). The NIHSS score was serially checked at baseline and at 1, 3, 7, and 14 days after admission. The hospital course was determined on day 7 after admission, and marked deterioration was defined as an increase in the NIHSS score of at least 4 during the first 7 days of hospitalization, as used previously.<sup>14</sup>

Two endpoints were evaluated at 6 months after ischemic stroke: poor outcome and excellent outcome. The BI and mRS score are the outcome measures that have been used most frequently in studies focusing on stroke-related disability and recovery of motor function after stroke. Both the mRS score and BI have been shown to be reliable and valid for use in stroke.<sup>15</sup> To define the outcome for stroke patients, we used a global endpoint that combined the mRS score and BI for the following reasons. First, this combination has been reported to be more powerful than the BI endpoints, and it reduces the required sample size.<sup>16</sup> Second, the mRS score may be less reproducible owing to its relative lack of structure,<sup>17</sup> whereas the BI has a U-shaped distribution in which patient outcomes cluster at the extremes.

The BI and mRS score were serially checked for more than 6 months (up to 1 year). We defined a poor outcome as any of the following endpoints: death, mRS score of >3, or BI of <60.<sup>18</sup> An excellent outcome was defined as reaching both of the following endpoints: mRS score of 0 or 1, and BI of  $\geq$ 95.

## 3. Statistical analysis

Receiver operating characteristic (ROC) curves were used to assess the usefulness of the individual NIHSS scores in predicting a poor or good outcome. We assessed discrimination by calculating the area under the ROC curve of sensitivity versus 1 minus the 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 best model for each outcome in each test cohort was defined as the model with the largest ROC curve area or, if there were no statistically significant differences in the areas, the model with the most practical curve. The sensitivity, specificity, and positive and negative predictive values of the clinical predictions - along with their associated 95% confidence intervals (CIs) - were calculated and plotted on the ROC curve to validate the individual scores in predicting poor and good outcomes at 6 months.

Stepwise logistic regression was used to assess which subset of variables best predicted a good or poor outcome as defined above. In addition to the NIHSS scores, several baseline variables were collected at enrollment: patient age and sex; presence of conventional vascular risk factors (hypertension, diabetes, hyperlipidemia, and atrial fibrillation); history of stroke or coronary heart disease; serum levels of inflammatory markers (C-reactive protein and fibrinogen); presence of metabolic syndrome; stroke subtypes (atherosclerotic, cardioembolic, cryptogenic, other determined etiology); vascular status (stenosis of  $\geq$ 50% or occlusion); DWI lesion volume; and presence of aggravating factors during the first 3 days of admission, such as admission hyperglycemia (blood sugar >300 mg/dl), hyperthermia (body temperature >38.5°C), hypoxia, or hypotension (systolic blood pressure <100 mmHg or sudden drop of

>30 mmHg). Most of the parameters were dichotomized (normal versus abnormal) in order to minimize the number of variables, improve the reliability of data collection, and simplify the clinical procedures. Those variables that were significant at the  $P<0.2$  level were entered into the initial multivariate model. When the most parsimonious model was obtained by backward stepwise elimination of the nonsignificant factors, each of the excluded variables was again entered separately into the model to test its contribution to the final model. The results are given as odds ratio (OR) estimates of relative risk, with the 95% CI. The Hosmer-Lemeshow goodness-of-fit chi-square test was used to assess the calibration of each model. This test compares the expected and observed distribution of cases and controls across deciles of predicted risk. Therefore, for this test, a higher probability value corresponds to a model with a better fit.

Statistical analysis was performed using the computer software SPSS (version 12.9 for Windows, SPSS, Chicago, Illinois, USA) and Medcalc (version 6.0 for Windows,). Statistical significance was established at the  $P<0.05$  level.

## RESULTS

### 1. Predictor and outcome characteristics

The patient characteristics of the predictor variables are listed in Table 1. With regard to the stroke mechanism, 253 patients (58%) were classified with atherosclerotic stroke, 97 (22%) with cardioembolic stroke, 67 (15%) with cryptogenic stroke, and 20 (5%) with other determined etiology. Systemic causes of aggravation were present in 59 patients (13.5%).

The baseline NIHSS scores are shown in Fig. 2. About 60% and 10% of the patients had scores of  $\leq 5$  and  $>13$ , respectively. There was a significant correlation between the DWI lesion volume and the baseline NIHSS score ( $r=0.588$ ,  $P<0.001$ ) (Fig. 2-B). The NIHSS score tended to decrease during the first 7 days of admission, although this was not statistically significant:  $5.56\pm 5.27$  at baseline,  $5.17\pm 5.17$  at day 1,  $4.77\pm 5.41$

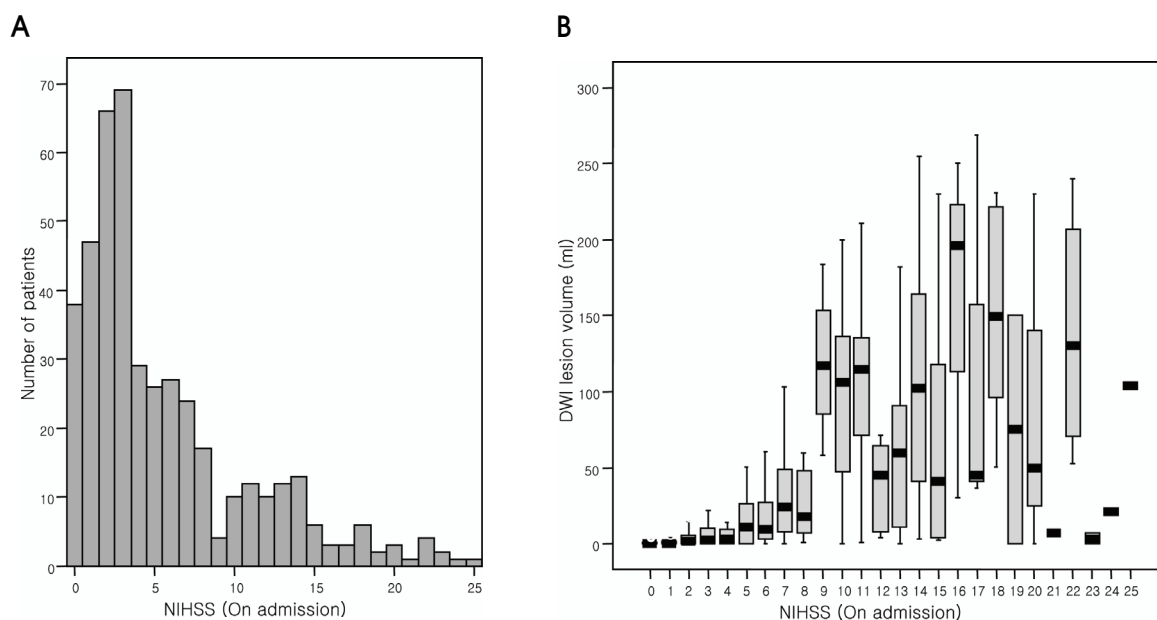
**Table 1.** Predictor-variable characteristics

Predictor variable	Characteristic
Age, year	62.1 $\pm$ 12.8
Sex	
Male	263 (60%)
Female	174 (40%)
Risk factors	
Hypertension	238 (54.7%)
Diabetes	116 (26.7%)
Atrial fibrillation	73 (16.7%)
Hyperlipidemia	80 (18.4%)
Stroke history	103 (23.7%)
CAOD history	46 (10.5%)
Metabolic syndrome	48 (32.4%)
C-reactive protein	0.64 $\pm$ 1.81
Fibrinogen	373.0 $\pm$ 125.7
Stroke subtype	
Atherosclerotic	253 (57.9%)
Cardioembolic	97 (22.2%)
Cryptogenic	67 (15.3%)
Other determined etiology	20 (4.6%)
DWI lesion volume, ml	30.56 $\pm$ 55.18
Vascular status	
Stenosis ( $\geq 50\%$ )	178 (41.5%)
Occlusion	143 (32.7%)
Systemic aggravating factors	
Hyperpyremia	24 (5.5%)
Hyperglycemia	7 (1.6%)
Hypoxia or hypotension	28 (6.4%)

at day 3, and  $4.38\pm 5.64$  at day 7. Moreover, marked changes ( $\geq 4$ ) in the NIHSS score were observed during this period, reflecting improvement in 59 patients (13.5%) and worsening in 24 patients (5.5%).

The patient outcomes are listed in Table 2. More than 50% of the patients showed excellent outcome, whereas about 25% showed poor outcome and 24 patients (5.5%) died.

### 2. Optimal time points and threshold for different endpoints



**Figure 2.** (A) Frequency distribution of baseline NIHSS scores in the study population. (B) Infarct volume as measured by DWI for each baseline NIHSS score in the study population.

**Table 2.** Outcome-variable characteristics

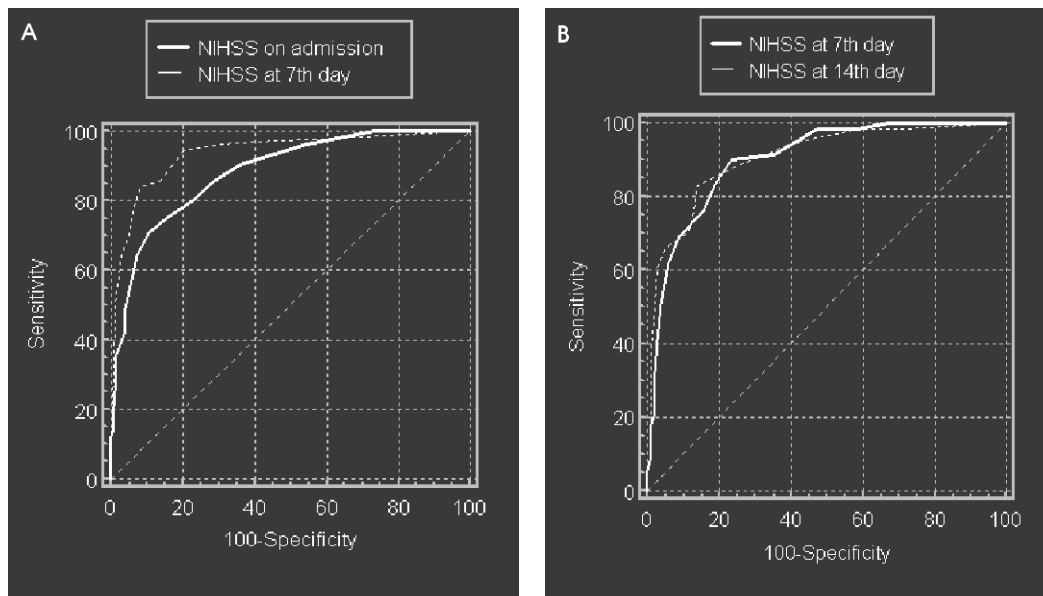
Outcome variable	7 days	3 month	6 months
<b>BI</b>			
≥95	210 (50.4%)	234 (53.5%)	254 (58.1%)
60 to <95	78 (18.7%)	95 (21.7%)	81 (18.5%)
<60	129 (30.9%)	108 (24.7%)	102 (23.3%)
<b>Modified Rankin scale score</b>			
0 or 1	188 (45.0%)	209 (47.8%)	237 (54.2%)
2 or 3	113 (27.0%)	131 (30.0%)	111 (25.4%)
4 or 5	117 (28.0%)	79 (18.1%)	65 (14.9%)
<b>Global endpoint</b>			
Poor outcome	133 (30.4%)	110 (25.2%)	108 (24.7%)
Excellent outcome	174 (39.8%)	199 (45.6%)	238 (54.5%)
<b>Death</b>			
	-	18 (4.1%)	24 (5.5%)

The area under the ROC curve for each NIHSS score is listed in Table 3, and the ROC curves for a poor outcome are shown in Fig. 3. The area under the ROC curve varied between 0.89 and 0.94 (with 0.5 indicating no discrimination and 1 indicating perfect discrimination). The discrimination was better for the NIHSS score determined 7 days after admission than for earlier NIHSS scores; the areas under the ROC curve for poor

outcome were 0.887, 0.927, 0.938, and 0.908 for the baseline, 3rd-, 7th-, and 14th-day NIHSS scores, respectively. Statistical comparisons suggested that the NIHSS score taken 7 days after admission had a better predictive performance than did the baseline NIHSS score ( $P=0.003$ , difference=0.054), whereas the area under the ROC curve did not differ significantly between the 7th- and 14th-day NIHSS scores ( $P=0.610$ ,

**Table 3.** Areas under the ROC curves

Outcome at 6 months:	Area under the ROC curve (95% CI)	
	Poor prognosis	Excellent prognosis
Baseline NIHSS score	0.887 (0.853-0.915)	0.872 (0.837-0.902)
3rd-day NIHSS score	0.927 (0.898-0.950)	0.894 (0.861-0.922)
7th-day NIHSS score	0.938 (0.911-0.959)	0.898 (0.866-0.925)
14th-day NIHSS score	0.908 (0.854-0.946)	0.927 (0.878-0.961)

**Figure 3.** Comparisons of the ROC curve between (A) baseline and 7th-day NIHSS scores, and (B) 7th- and 14th-day NIHSS scores.

difference=0.006) (Fig. 3). For an excellent prognosis, there was no significant difference between the baseline and 7th-day NIHSS scores ( $P=0.104$ , difference=0.025) or between the 7th- and 14th-day NIHSS scores ( $P=0.464$ , difference=0.007).

The NIHSS score strongly predicted the likelihood of patient recovery after stroke. An NIHSS score of  $\geq 6$  at 7 days after admission predicted a poor outcome with a sensitivity of 84% (95% CI, 76-90%), a specificity of 92% (95% CI, 88-94%), and positive and negative predictive values of 77% and 95%, respectively. The best predictor for an excellent prognosis at 7 days was an NIHSS score of  $\leq 3$ , which gave a sensitivity of 91% (95% CI, 87-94%), a specificity of 75% (95% CI, 68-81%), and positive and negative predictive values of 81% and 87%, respectively.

### 3. Prognostic models for long-term outcomes

Table 4 lists the OR and CI values for the factors in the multiple logistic regression model. The significant factors in the model were similar between the two prognostic endpoints. Age, DWI lesion volume, and the NIHSS score at 7 days after admission were independently associated with both prognostic endpoints, with the last being the strongest predictor. Stroke history was an independent predictor for poor outcome - but not for excellent outcome - at 6 months after symptom onset. Patients with an NIHSS score of  $\geq 6$  at 7 days after admission were about 52 times more likely to remain in a severely disabled state at 6 months after stroke than were patients with an NIHSS score of  $< 6$ , after adjusting

**Table 4.** Multiple logistic regression test

Variable	Poor outcome		Excellent outcome	
	OR (95% CI)	<i>P</i>	OR (95% CI)	<i>P</i>
Age, per 1-year increase	1.15 (1.04-1.27)	0.005	0.90 (0.90-0.99)	0.018
Stroke history	11.24 (1.25-100.96)	0.031	N/A	0.321
DWI lesion volume, per 1-ml increase	1.03 (1.01-1.15)	0.012	0.98 (0.96-0.99)	0.038
NIHSS score at day 7				
≥6	51.97 (6.57-411.37)	<0.001	-	-
≤3	-	-	17.427 (5.291-57.40)	<0.001

**Table 5.** Differences between predictor models

Reference	No. of patients	Enrollment and endpoint	Factors analyzed	Predictor models
Adams HP, et al. <sup>12</sup>	1,281	Any stroke within 24 hours Global endpoint at 3 months	Baseline NIHSS score, age/sex, stroke subtype, stroke history	Baseline NIHSS score, lacunar stroke, age, sex, race, previous stroke
Counsell C, et al. <sup>19</sup>	530	Any first-ever stroke within 30 days Survival at 30 days and independent state at 6 months	Age/sex, social status, prestroke disability, risk factors, Glasgow Coma Scale score, neurological deficits, laboratory results	Six simple variables (age, living alone, independence in ADL before stroke, verbal component of Glasgow Coma Scale, arm power, ability to walk)
Muir KW, et al. <sup>4</sup>	373	Any acute stroke Alive at home or in care, or dead at 3 months	Baseline NIHSS score, Guy's prognostic score, The Canadian Neurological Scale score, MCA Neurological Score	Baseline NIHSS score best predicts 3-month outcome. No other score added useful information to the NIHSS score. (All patients with baseline NIHSS score of (13 had poor outcome)
Tilling K, et al. <sup>20</sup>	299	Any stroke BI at 1 year	Age/sex, neurological deficits	Urinary incontinence, sex, prestroke disability, dysarthria
Weimar C, et al. <sup>6</sup>	1,023	Any stroke within 6 hours of onset BI at 100 days	Age/sex, stroke history, baseline NIHSS score, atrial fibrillation	Age, baseline NIHSS score
Baird AE, et al. <sup>21</sup>	66	Any acute stroke within 48 hours of onset BI at 1 to 3 months	Age/sex, history of hypertension or heart disease, NIHSS score, entry into drug trial, time from onset of symptoms to magnetic resonance imaging, DWI lesion volume	Combined measurements of the NIHSS score, time in hours from stroke onset to DWI, volume of ischemic brain tissue on DWI gave the best prediction of stroke recovery
This study	437	Nonlacunar stroke in MCA territory within 72 hours Global endpoints at 6 months	Several additional factors, including the DWI lesion volume, vascular status, inflammatory markers, metabolic syndrome markers, systemic causes of deterioration	NIHSS score at day 7, age, previous stroke, DWI lesion volume

for other factors. Similarly, patients with an NIHSS score of  $\leq 3$  at 7 days after admission were about 17 times more likely to regain independent life, after adjusting for other factors. The multilevel model was internally valid, as shown by the good fit of the model on the study population for a poor prognosis ( $\chi^2=10.497$ ,  $df=8$ ,  $P=0.232$ ) and an excellent prognosis ( $\chi^2=10.497$ ,  $df=8$ ,  $P=0.232$ ).

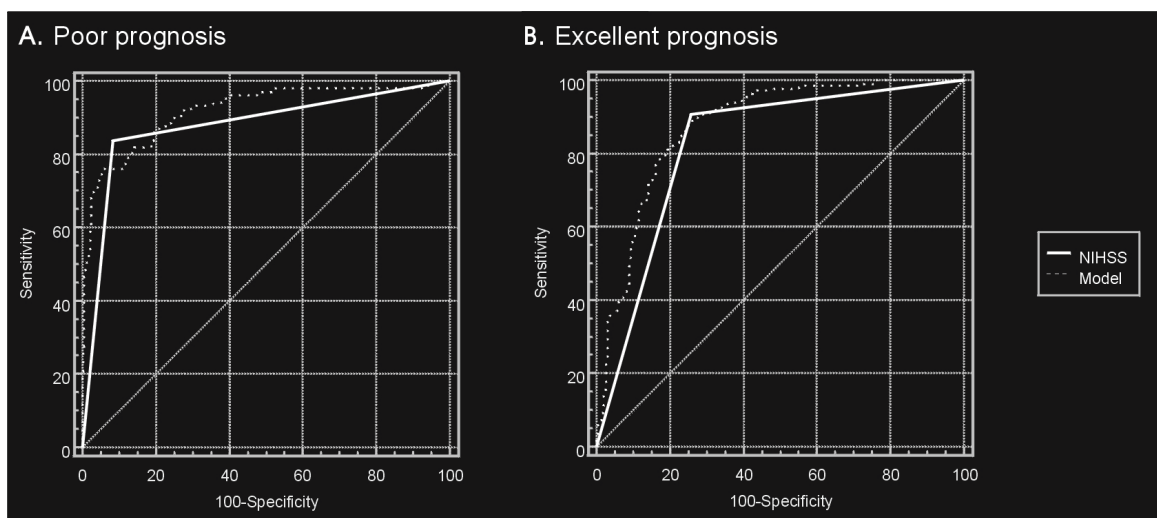
To develop prognostic models for long-term outcome, point values were assigned to each factor by multiplying the  $\beta$  coefficients from the logistic regression model by 33 for a poor prognosis and 50 for an excellent prognosis, and rounding off to the nearest integer. The resulting point values assigned to each factor used in calculating the poor-prognosis risk index were (a) 1 point for each milliliter infarct volume, (b) 4.7 points for each year of age, (c) 80 points for the presence of previous stroke, and (d) 130 points for the presence of severe deficits at 7 days after admission, defined as an NIHSS score of  $\geq 6$ . Similarly, the points for an excellent-prognosis risk index were (a) -1 point for each milliliter of infarct volume, (b) -2.8 points for each year of age, and (c) 143 points for the presence of mild deficits at 7 days after admission, defined as an NIHSS score of  $\leq 3$ .

The ROC curves for the NIHSS criteria and the models for poor and excellent prognoses are shown in

Fig. 4. We measured the areas under the ROC curves for poor and excellent prognoses using the NIHSS criteria and using a model that included age, DWI lesion volume, and stroke history. For a poor prognosis at 6 months after stroke, the area under the ROC curve was 0.878 (95% CI, 0.842-0.907) using only the NIHSS criteria (score of  $\geq 6$  at 7 days after admission), and increased to 0.919 (95% CI, 0.888-0.943) using the prognostic model that included age, DWI lesion volume, and stroke history. However, statistical comparisons indicated that the area under the ROC curve did not differ significantly between the NIHSS criteria and the prognostic model ( $P=0.079$ , difference=0.041). Discriminating an excellent prognosis was better for the prognostic model than for the NIHSS criteria (score of  $\leq 3$  at 7 days after admission). The areas under the ROC curves for an excellent outcome were significantly higher for the combined prognostic model than for the NIHSS criteria (0.876 vs 0.825, respectively;  $P=0.004$ ).

## DISCUSSION

Our study is unique in that (a) our patients represented a homogeneous group (those with nonlacunar stroke within the MCA territory) due to our exclusion of patients with lacunar stroke and posterior circulation



**Figure 4.** Comparisons of the ROC curve between the NIHSS criteria and the models for (A) a poor prognosis and (B) an excellent prognosis.



stroke, and (b) DWI data were analyzed. Most previous studies analyzed stroke of several etiologies and considered different factors (Table 5).<sup>4,6,12,19-21</sup> Patients with lacunar stroke are known to have an excellent long-term prognosis regardless of their baseline clinical characteristics,<sup>12</sup> and are therefore not suitable candidates for most acute- or subacute-stroke trials. For patients with posterior circulation stroke, it has been reported that there is no significant correlation between the infarct size as measured by the DWI volume and the NIHSS score.<sup>22</sup> Indeed, we found that the predictive performance of the serial NIHSS scores was worse in 187 patients with posterior circulation stroke (data not shown) than in the patients with MCA territory infarcts reported here.

The accurate prediction of outcome in the acute and subacute phases of stroke would be of value in both epidemiological research and clinical practice.<sup>19</sup> Although there have been numerous investigations of prognostic models to predict long-term stroke outcome, most of these have focused on patients in the hyperacute stage of stroke. The present study has yielded a prognostic model for patients with subacute stroke that may guide patient management in clinical practice or patient enrollment in clinical trials. Our results of prognostic factors in subacute stroke were different from those in acute-stroke trials. The model for acute-stroke trials could not be used in the subacute-stroke trial for several reasons. The results of the present study show that the major prognostic factors at the subacute stage of stroke were ischemic injury (DWI lesion volume), the ability to recover functional level (patient age), or both (NIHSS score at the subacute stage). In an acute-stroke trial, stroke outcome can be greatly influenced by numerous factors, such as recanalization, evolving stroke, reperfusion injury, and collateral circulation. However, the influence of these factors on stroke outcome would be lower in a subacute-stroke trial than in an acute-stroke trial. Our logistic regression analysis of prognostic factors at 7 days after admission in the present study showed that systemic factors of aggravation (e.g., initial hyperglycemia or hyperpyremia) and degree of stenosis were not independent predictors of long-term outcome. Our results are consistent with those of a previous study

in which a significant number of patients with acute ischemic stroke treated with conventional therapy showed early improvement or worsening, as assessed by the NIHSS score.<sup>23</sup> These results indicate that the NIHSS scores at a more optimal time point - relative to baseline - should be used in evaluations for subacute-stroke trials.

In the present study, the NIHSS score at 7 days after admission was the best at predicting the 6-month outcome. Patients with severe neurological deficits after acute ischemic stroke, as quantified by the NIHSS, have poor prognoses. We found the discriminatory power in predicting stroke outcome was higher for the NIHSS score taken at 7 days after admission than for the baseline NIHSS score, suggesting that early changes in stroke scores influence outcome predictions. After the 1st week following admission, it is possible to identify a subset of patients who are highly likely to experience a poor outcome. Our later NIHSS score (i.e., that at day 14) did not provide useful additional information in this study. It has been shown that the clinical course of recovery stabilizes beyond day 4, with improvement occurring linearly over time from then on.<sup>24</sup> Cutoff NIHSS scores of 6 and 3 were the best at predicting a poor prognosis and an excellent prognosis at 6 months after stroke, respectively.

Our results have important implications for the choice of patients enrolled in subacute-stroke trials. Spontaneous recovery occurs frequently in patients with a good prognosis after stroke. Conversely, patients with very severe stroke may be destined for a poor outcome, regardless of any intervention.<sup>25</sup> The NIHSS score can be used as an exclusion or inclusion criterion for the enrollment of patients in trials testing new treatments for stroke, as follows. For subacute-stroke trials that may be harmful, patient enrollment should be restricted to those with an NIHSS score of  $\geq 6$  at 7 days after admission; enrollment in a trial that presents significant risk should require an even higher score. Similarly, patients with an NIHSS score of  $\leq 3$  should not be included in stroke trials, even if the treatment modality is not extensive or hazardous. The recruitment of patients with low NIHSS scores into a trial designed to test a promising intervention would likely obscure any treatment effect and

thus could increase the likelihood of rejecting a potentially beneficial therapy.

In addition to the NIHSS score after stroke, differences in baseline patient characteristics can also influence study group outcomes. In previous studies, several clinical characteristics have been shown to influence the final clinical status: for example, age,<sup>26</sup> stroke severity,<sup>12</sup> and stroke subtype<sup>27</sup> each influence the mortality rate. Our logistic regression analysis revealed that age, presence of previous stroke, DWI lesion volume, and NIHSS score at 7 days after admission were independent predictors of the long-term outcome. Our results are in agreement with those of other studies performed hyperacute-stroke settings.<sup>21</sup> However, we have also shown that no factor used in addition to the 7th-day NIHSS score improved the prediction of a poor long-term outcome. For excellent prognosis, the discrimination was marginally better with the prognostic model that included the other factors than with the NIHSS criteria alone.

This study has several limitations. Our population was a hospital-based cohort, and unselected patients in different care settings might have prognoses that differ from those suggested in our models. In other words, our results can only be considered valid for patients in Korean acute-stroke units and cannot be automatically extrapolated to stroke registers or other stroke-care institutions. Although the stroke history was analyzed in our study, prestroke disability was not considered. Several studies have shown that prestroke disability affects the level of outcome after stroke,<sup>19,20</sup> and hence further studies involving patients with first-time strokes are needed.

In summary, our study shows that the NIHSS score at 7 days after admission provides a clinically useful discrimination point for accurate forecasting of long-term outcomes after stroke. We issue a caution that randomization into clinical trials without stratification of stroke severity increases the risk of testing two patient populations with different clinical courses. Our data may be helpful in predicting long-term prognosis as well as in decision-making concerning novel therapeutic applications in the subacute stage after stroke.

## ACKNOWLEDGMENTS

This work was supported by a grant from the Korea Health 21 R&D Project, Ministry of Health and Welfare, Republic of Korea (no. 0412-DB00-0101-007) to O.Y.B. We thank Professor Seung Soo Sheen for his advice and assistance in preparing the manuscript. The study would not have been possible without the help of the patients and their families, and the generous cooperation of doctors and nurses at Ajou University Hospital.

## REFERENCES

1. De Haan R, Horn J, Limburg M, Van Der Meulen J, Bossuyt P. A comparison of five stroke scales with measures of disability, handicap, and quality of life. *Stroke* 1993;24:1178-1181.
2. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 1995; 333:1581-1587.
3. Fiorelli M, Alperovitch A, Argentino C, Sacchetti ML, Toni D, Sette G, et al. Prediction of long-term outcome in the early hours following acute ischemic stroke. Italian Acute Stroke Study Group. *Arch Neurol* 1995;52:250-255.
4. Muir KW, Weir CJ, Murray GD, Povey C, Lees KR. Comparison of neurological scales and scoring systems for acute stroke prognosis. *Stroke* 1996;27:1817-1820.
5. Jorgensen HS, Nakayama H, Raaschou HO, Olsen TS. Acute stroke. Prognosis and a prediction of the effect of medical treatment on outcome and health care utilization. The Copenhagen Stroke Study. *Neurology* 1997;49:1335-1342.
6. Weimar C, Konig IR, Kraywinkel K, Ziegler A, Diener HC; German Stroke Study Collaboration. Age and National Institutes of Health Stroke Scale Score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. *Stroke* 2004;35:158-162.
7. Braitman LE, Davidoff F. Predicting clinical states in individual patients. *Ann Intern Med* 1996;125:406-412.
8. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-387.
9. Kondziolka D, Wechsler L, Goldstein S, Meltzer C, Thulborn KR, Gebel J, Jannetta P, et al. Transplantation

- of cultured human neuronal cells for patients with stroke. *Neurology* 2000;55:565-569.
10. Bang OY, Lee JS, Lee PH, Lee G. Autologous mesenchymal stem cell transplantation in stroke patients. *Ann Neurol* 2005;57:874-882.
  11. Kwakkel G, Wagenaar RC, Kollen BJ, Lankhorst GJ. Predicting disability in stroke - a critical review of the literature. *Age Ageing* 1996;25:479-489.
  12. Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, et al. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). *Neurology* 1999;53:126-131.
  13. Bang OY, Lee PH, Joo SY, Lee JS, Joo IS, Huh K. Frequency and mechanisms of stroke recurrence after cryptogenic stroke. *Ann Neurol* 2003;54:227-234.
  14. Brott TG, Haley EC Jr, Levy DE, Barsan W, Broderick J, Sheppard GL, et al. Urgent therapy for stroke. Part I. Pilot study of tissue plasminogen activator administered within 90 minutes. *Stroke* 1992;23:632-640.
  15. Duncan PW, Jorgensen HS, Wade DT. Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice. *Stroke* 2000;31:1429-1438.
  16. Young FB, Lees KR, Weir CJ. Glycine Antagonist in Neuroprotection International Trial Steering Committee and Investigators. Strengthening acute stroke trials through optimal use of disability end points. *Stroke* 2003;34:2676-2680.
  17. Wilson JT, Hareendran A, Grant M, Baird T, Schulz UG, Muir KW, et al. Improving the assessment of outcomes in stroke: use of a structured interview to assign grades on the modified Rankin Scale. *Stroke* 2002;33:2243-2246.
  18. Sulter G, Steen C, De Keyser J. Use of the Barthel index and modified Rankin scale in acute stroke trials. *Stroke* 1999;30:1538-1541.
  19. Counsell C, Dennis M, McDowall M, Warlow C. Predicting outcome after acute and subacute stroke: development and validation of new prognostic models. *Stroke* 2002;33:1041-1047.
  20. Tilling K, Sterne JA, Rudd AG, Glass TA, Wityk RJ, Wolfe CD. A new method for predicting recovery after stroke. *Stroke* 2001;32:2867-2873.
  21. Baird AE, Dambrosia J, Janket S, Eichbaum Q, Chaves C, Silver B, et al. A three-item scale for the early prediction of stroke recovery. *Lancet* 2001; 357:2095-2099.
  22. Linfante I, Llinas RH, Schlaug G, Chaves C, Warach S, Caplan LR. Diffusion-weighted imaging and National Institutes of Health Stroke Scale in the acute phase of posterior-circulation stroke. *Arch Neurol* 2001;58:621-628.
  23. Wityk RJ, Pessin MS, Kaplan RF, Caplan LR. Serial assessment of acute stroke using the NIH Stroke Scale. *Stroke* 1994;25:362-365.
  24. Toni D, Fiorelli M, Bastianello S, Falcou A, Sette G, Ceschin V, et al. Acute ischemic strokes improving during the first 48 hours of onset: predictability, outcome, and possible mechanisms. A comparison with early deteriorating strokes. *Stroke* 1997; 28:10-14.
  25. Dorman PJ, Sandercock PA. Considerations in the design of clinical trials of neuroprotective therapy in acute stroke. *Stroke* 1996;27:1507-1515.
  26. Petty GW, Brown RD Jr, Whisnant JP, Sicks JD, O'Fallon WM, Wiebers DO. Survival and recurrence after first cerebral infarction: a population-based study in Rochester, Minnesota, 1975 through 1989. *Neurology* 1998;50:208-216.
  27. Petty GW, Brown RD Jr, Whisnant JP, Sicks JD, O'Fallon WM, Wiebers DO. Ischemic stroke subtypes: a population-based study of functional outcome, survival, and recurrence. *Stroke* 2000;31:1062-1068.