Prediction of Functional Outcome After Stroke
Comparison of the Orpington Prognostic Scale and the NIH Stroke Scale

Sue-Min Lai, PhD, MS, MBA; Pamela W. Duncan, PhD, PT; John Keighley, MS

Background and Purpose—This study compared the ability of 2 stroke impairment scales, Orpington Prognostic Scale and National Institutes of Health (NIH) Stroke Scale, to predict disability as measured by the Barthel activities of daily living (ADL) Index and higher level of self-reported physical functioning as measured by the SF-36 physical functioning index (PFI) at 1, 3, and 6 months after stroke.

Methods—The participants in this ongoing study are 184 individuals who sustained an eligible stroke and were recruited for the Kansas City Stroke Study. All patients were prospectively evaluated using standardized assessments at enrollment (within 14 days of stroke onset) and followed at 1, 3, and 6 months after stroke. Coefficient of determination ($R^2$) was used to assess the ability of the 2 stroke scales to prognosticate outcomes.

Results—Means and SDs of the Orpington Prognostic Scale and NIH Stroke Scale measured at baseline were $3.6 \pm 1.31$ and $5.5 \pm 4.58$, respectively. The Spearman’s rank correlation between the 2 baseline measures was $0.83 (P=0.0001)$. The Orpington Prognostic Scale and the NIH Stroke Scale explained well the variance in Barthel ADL Index ($P<0.001$). However, the Orpington Prognostic Scale explained more variance than did the NIH Stroke Scale. Similarly, the Orpington Prognostic Score explained more variance in higher level of physical function than did the NIH Stroke Scale. The amount of variance in Barthel ADL Index and SF-36 PFI, which were explained by both stroke severity measures, decreased over time.

Conclusions—Our results demonstrate that in a sample of mostly mild and moderate strokes, the Orpington Prognostic Scale compared with the NIH Stroke Scale is simpler to use and is a slightly better predictor of ADL and higher levels of physical function. (Stroke. 1998;29:1838-1842.)

Key Words: activities of daily living ■ impairment ■ physical function ■ stroke

Stroke is heterogeneous in type and severity. To characterize probabilities of outcomes and plan for discharge, we need a stroke scale that is able to ascertain the precise nature of stroke-related impairment and to characterize severity. A good stroke scale identifies neurological impairments and is quantified so that the patient’s progress can be objectively monitored. It should provide a logical basis for treatment and predict future functional outcomes. Previous researchers have demonstrated that impairments are strongly associated with functional outcomes, but they only partially explain stroke-related disability.1-4 Nevertheless, a baseline stroke impairment scale can be used to assess stroke severity and to adequately predict functional outcome.

Several impairment scales are available for clinical practice and research.1-3 The National Institutes of Health (NIH) Stroke Scale is probably the most frequently used measure of stroke impairment.5-8 This stroke-specific scale has been widely used in clinical trials to measure baseline severity or progress associated with investigational therapies.5-8 A recent study by Muir et al9 shows that baseline NIH Stroke Scale predicts 3-month outcomes (alive at home, alive in care, or dead). Individuals who scored greater than 13 on the NIH Stroke Scale had very poor functional outcomes (alive in care or death) compared with those who scored 13 or less. The Orpington Prognostic Scale,10 which is modified from the Edinburgh score,11 is a simple stroke impairment scale, but it is not as well known or as commonly used as the NIH Stroke Scale. When assessed at 2 weeks after stroke, the Orpington score was shown by Kalra and Crome10 to be a useful indicator for 14-week poststroke activities of daily living (ADL) scores and discharge disposition. Both the Orpington Prognostic Scale and the NIH Stroke Scale appear to have good predictive validity but they require different skills and amounts of time for administration. The purpose of this study was to compare the ability of these two stroke impairment scales to predict disability as measured by the Barthel ADL Index12 and higher levels of self-reported physical functioning index (PFI) as measured by the SF-3613 at 1, 3, and 6 months after stroke.

Received April 2, 1998; final revision received May 29, 1998; accepted June 10, 1998.
From the Department of Preventive Medicine (S.-M.L., J.K.), and the Center on Aging (P.W.D.), University of Kansas Medical Center, Kansas City; the Department of Veteran Affairs Medical Center (P.W.D.), Kansas City, Mo; and the Department of Health Services Research, University of Kansas, Lawrence.
Correspondence to Sue-Min Lai, PhD, MS, MBA, Department of Preventive Medicine, University of Kansas Medical Center, 3901 Rainbow Blvd, Kansas City, KS 66160-7313. E-mail smlai@kumc.edu
© 1998 American Heart Association, Inc.
Subjects and Methods

The participants in this study are 184 individuals who sustained an eligible stroke and were recruited for the Kansas City Stroke Study. Case ascertainment for the Kansas City Stroke Study started in October of 1995 and the follow-up effort is ongoing. The eligible study participants were recruited from any of 12 participating hospitals in the Greater Kansas City area. Eligible stroke patients were identified by (1) a review of daily admission records, (2) referrals from physicians, clinical nurse specialists, and therapists on medical, neurology, and rehabilitation units, and (3) review of discharge codes. To be accepted into this study, the subject had to have a confirmed eligible stroke as defined by World Health Organization (WHO) criteria. The stroke was confirmed by clinical assessment and/or by a CT/MRI scan. A stroke was defined according to the WHO criteria as "rapid onset and of vascular origin reflecting a focal disturbance of cerebral function, excluding isolated impairments of higher function and persisting longer than 24 hours." Trained nurses/physical therapists reviewed medical records and interviewed both patients and physicians to determine whether the patient was eligible and consented for enrollment. Subjects were excluded if they (1) were less than 18 years of age; (2) had stroke onset more than 14 days earlier; (3) had stroke due to subarachnoid hemorrhage; (4) had hepatic failure; (5) had renal failure; (6) had New York Heart Association functional grade III/IV heart failure (ie, patients with cardiac disease resulting in inability or marked limitation to carry on any physical activity without discomfort); (7) were not expected to live 6 months; (8) lived in a nursing home prior to stroke; (9) were unable to take care of own affairs prior to stroke; (10) were lethargic, obtunded, or comatose; and (11) lived more than 70 miles from the participating hospital.

The patients were evaluated using a variety of standardized assessments at enrollment and followed at 1, 3, and 6 months after stroke by a study nurse/physical therapist at home or at a chronic care facility. Each study nurse/physical therapist received at least 2 weeks of training in the administration of the measures. All study nurses and physical therapists received certification in the administration of NIH Stroke Scale. Assessments included baseline demographics, stroke characteristics, Orpington Prognostic Scale, NIH Stroke Scale, Barthel ADL Index, and assessment of Prior Function on the physical domain of the SF-36 and SF-36 PFI. Measurements were performed at baseline (within 14 days of stroke onset), 1 month, 3 months, and 6 months after stroke. For the present study, only baseline measures from the Orpington Prognostic Scale and NIH Stroke Scale and follow-up measures from the Barthel ADL Index and SF-36 PFI were included in the analysis.

The Orpington Prognostic Scale and the NIH Stroke Scale were used to measure stroke severity at baseline. The Barthel ADL Index was used to measure basic activities of daily living, and SF-36 PFI measured the patients' higher level of physical functioning. The Barthel ADL Index includes measures of basic ADL including personal hygiene, bathing, feeding, toilet, stair climbing, dressing, bowel and bladder control, ambulation, and bed/chair transfers. The Barthel ADL Index scores range from 0 to 100, with 100 indicating patient is fully independent in physical functioning. The SF-36 includes 8 domains. The present analysis only included SF-36 PFI, which measures higher level of physical functioning (vigorous and moderate activities, lifting or carrying groceries, climbing several flights of stairs, climbing 1 flight of stairs, bending, kneeling or stooping, walking more than a mile, walking several blocks, walking 1 block, and bathing or dressing). The score for the SF-36 PFI ranges from 0 to 100, with 100 indicating patient is fully independent.

Descriptive statistics were used to show demographics, prior functional status, stroke characteristics, severity of impairment due to stroke, and scores on the Barthel Index and the SF-36 PFI measured at 1, 3, and 6 months after stroke. Correlation between the Orpington Prognostic Scale and the NIH Stroke Scale was calculated by the Spearman rank correlation. Although the sum of item scores (eg, Barthel Index) is ordinal in nature, the prognostic ability of the Orpington Prognostic Scale compared with that of the NIH Stroke Scale was examined with linear regression analyses. However, it has been argued that in some cases ordinal-level data may be treated as interval-level data without serious problem. The R² and adjusted R² values were used to measure the extent to which an outcome can be explained by a stroke severity scale. The explanatory ability of each of the domains of the Orpington Prognostic Scale and the NIH Stroke Scale was further analyzed using partial R² from linear regression analysis with a forward selection procedure.

### TABLE 1. Orpington Prognostic Scale

<table>
<thead>
<tr>
<th>A. Motor deficit in arm</th>
<th>Lying supine, patient flexes shoulder to 90° and is given resistance.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0=MRC grade 5 (normal power)</td>
<td>0.4=MRC grade 4 (diminished power)</td>
</tr>
<tr>
<td>0.8=MRC grade 3 (movement against gravity)</td>
<td>1.2=MRC grade 1–2 (movement with gravity eliminated or trace)</td>
</tr>
<tr>
<td>1.6=MRC grade 0 (no movement)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Proprioception (eyes closed)</th>
<th>Locates affected thumb:</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0=Accurately</td>
<td>0.4=Slight difficulty</td>
</tr>
<tr>
<td>0.8=Feeds thumb via arm</td>
<td>1.2=Unable to find thumb</td>
</tr>
</tbody>
</table>

### C. Balance

<table>
<thead>
<tr>
<th>0.0=Walks 10 feet without help</th>
<th>0.4=Maintains standing position (unsupported for 1 min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8=Maintains sitting position</td>
<td>1.2=No sitting balance</td>
</tr>
</tbody>
</table>

### D. Cognition

Hodkinson’s Mental Test: Score one point for each correct answer.

1. **1. Age of patient**
   - 2. **Time (to the nearest hour)**
   - I am going to give you an address, please remember it and I will ask you later: 42 West Street.
   - 3. **Name of hospital**
   - 4. **Year**
   - 5. **Date of birth of patient**
   - 6. **Month**
   - 7. **Years of the Second World War**
   - 8. **Name of the President**
   - 9. **Count backwards (20-1)**
   - 10. **What is the address I asked you to remember?**
   - 42 West Street
   - 0.0=Mental test score of 10
   - 0.4=Mental test score of 8–9
   - 0.8=Mental test score of 5–7
   - 1.2=Mental test score of 0–4

**TOTAL SCORE:** 1.6=Motor + Proprioception + Balance + Cognition
Results

One hundred and eighty-four subjects enrolled in the Kansas City Stroke Study were included in the present analysis. All subjects were community dwelling before their strokes. Subject demographics and associated stroke characteristics are shown in Table 2. By the end of the 6 months, 9 patients had died, 12 refused to be in study, 3 were withdrawn by a family member, 2 moved out of the study area, and 1 was unable to schedule an appointment because of family problems.

Descriptive statistics on baseline stroke severity measured within 14 days after stroke (mean ± SD, range, 0 to 14 days; range, 0 to 14 days) and subjects’ prior functional scores are shown in Table 2. Assessments on 48% of the patients were made during the first week of stroke onset. Descriptive statistics on scores of Barthel ADL Index and SF-36 PFI measured at 1 month, 3 months, and 6 months after stroke are shown in Table 3. The Orpington Prognostic Score and the NIH Stroke Scale baseline measures were correlated (r = 0.83, P < 0.0001).

Table 4 summarizes the results of multiple linear regression analysis for the Orpington Prognostic Scale and the NIH Stroke Scale. Both the Orpington Prognostic Scale and the NIH Stroke Scale explained well the variance in Barthel ADL Index (P < 0.001). However, the Orpington Prognostic Scale explained more variance than did the NIH Stroke Scale (Table 4). The amount of variance explained by both measures decreased over time.

The variances in higher level function (SF-36 PFI) explained by the baseline Orpington Prognostic Score and NIH Stroke Scale are also shown in Table 4. The Orpington Prognostic Score explained more variance in higher level of physical function than did the NIH Stroke Scale. The amount of variance explained by both measures decreased over time. Neither the Orpington Prognostic Score nor the NIH Stroke Scale explained as much of the variance of higher level physical functioning as they did for basic ADL.

The explanatory ability of the Orpington Prognostic Score for the Barthel ADL Index at 1 month was primarily from balance (51%) followed by motor deficit in arm (6%), cognition (6%), and proprioception (1%). (Table 5) The percentages of variability from the SF-36 PFI at 1 month, explained by Orpington Prognostic Score, were 34% from balance; 3% from motor deficit in arm, and 0.4% from cognition. When each domain of the NIH Stroke Scale was examined, arm strength was shown to contribute the most in predicting Barthel ADL (48%), followed by leg strength (7%), level of consciousness (2%), and sensory (1%). The domains and percentages of variances in SF-36 PFI explained by each corresponding NIH Stroke Scale domain (Table 5) were arm strength (27%), leg (2%), consciousness (1%), language (1%), and vision (1%).

Discussion

The Orpington Prognostic Scale and the NIH Stroke Scale are impairment level measures that are strongly correlated. However, the Orpington Prognostic Scale has a slightly higher predictive ability compared with that of the NIH Stroke Scale. The Orpington Prognostic Score explained more of the variance in basic ADL and higher level physical functions at 1, 3, and 6 months after stroke. When the level of impairment on the Orpington Prognostic Scale was categorized as minor, moderate, or major, the predictive ability on functional outcome remained. The predictive ability of the NIH Stroke Scale with 2 severity levels (NIH score > 13 versus NIH score ≤ 13) was lower than the Orpington Prognostic Score.
13) declined, although this analysis was based on a small number of cases with NIH score ≤13 (n=13).

The NIH Stroke Scale cut point of 13 was selected on the basis of the previous research of Muir et al. Muir and colleagues reported that with a cut point of 13 the NIH Stroke Scale predicted 3-month poor functional outcome with a 0.71 sensitivity, a specificity of 0.90, and an overall accuracy of 0.83. In their study, alive in care or death at 3 months after stroke was used as a marker for poor 3-month functional outcome. The functional outcome that did not separate alive in care from death may be a reason why the NIH Stroke Scale with a cut point of 13 may be very predictive of global outcomes but not as useful in the prediction of functional outcomes (Barthel basic ADL and higher levels of physical function). In their analysis, other neurological scales (Canadian Neurological Scale, Middle Cerebral Artery Neurological Score, and the Guy’s Prognostic Score) were shown to have predictive values similar to that of the NIH Stroke Scale for global outcomes. Our analysis of ADL and higher physical function outcomes demonstrated that a cutoff point of 13 for the NIH Stroke Scale has low predictive value. Our outcome analysis was the prospective assessment of ADL and higher physical function instead of global outcomes, which were acquired by record linkage to death records and hospital discharge records. Subsequently, our results used different outcome measures that did not support the value of using 13 as the cutoff for the NIH Stroke Scale.

In a cross-sectional study of individuals 6 months after stroke, De Haan and colleagues reported that the variance in Barthel ADL Index explained by impairment level measures including the NIH Stroke Scale was less than 50% ($R^2=47.50\%$). They also reported that the relationships between impairment level measures and more distal health status measures like the Sickness Impact Profile were weak ($R^2=33\%$). Our prospective assessment of the relationship between baseline impairment and functional outcome at various times showed similar trends. In other words, our study results showed that impairments are more strongly associated with ADL than other measures of health status and that the variance explained by both measures decreased over time. This is expected because disability limitations may be minimized over time by development of compensatory strategies and good physical, environmental, and social support. Other factors such as depression and urinary incontinence may also modify levels of physical function.

Analysis of the domains of the Orpington Prognostic Score and the NIH Stroke Scale, which explained the variance in outcomes, revealed that arm power in the NIH Stroke Scale explained most (48\%) of the variance in ADL outcomes, whereas in the Orpington Prognostic Scale balance explained 51\% of the variance in ADL outcomes. Balance is a fundamental component of physical functioning and was incorporated in the Orpington Prognostic Scale and not in the NIH Stroke Scale. This may explain a slightly better prediction of physical function by stroke severity when the Orpington Prognostic Scale was used.

The Orpington Prognostic Scale is easy to use, requires less than 5 minutes to perform the test, and requires no extensive training, whereas the NIH Stroke Scale requires extensive

### Table 4. Prognostic Ability of the Orpington Prognostic Scale and the NIH Stroke Scale

<table>
<thead>
<tr>
<th>Stroke Scale</th>
<th>Barthel Index</th>
<th>SF-36 PFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke scales were analyzed as continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPS</td>
<td>62</td>
<td>30</td>
</tr>
<tr>
<td>NIHSS</td>
<td>56</td>
<td>22</td>
</tr>
<tr>
<td>Stroke scales were analyzed as categorical*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPS</td>
<td>61</td>
<td>26</td>
</tr>
<tr>
<td>NIHSS</td>
<td>32</td>
<td>4</td>
</tr>
</tbody>
</table>

SF-36 PFI indicates short form 36 physical functioning index; OPS, Orpington Prognostic Scale; NIHSS, National Institutes of Health Stroke Scale. All $R^2$ or adjusted $R^2$ have a $P<0.001$.

*OPS has been categorized as minor (OPS <3.2), moderate (3.2≤OPS≤5.2), or major (OPS >5.2). NIH Stroke Scale has been categorized as ≤13 or >13.

### Table 5. Linear Regression Models to Explain Functional Outcome at 1 Month by Domains of Orpington Prognostic Scale and NIH Stroke Scale

<table>
<thead>
<tr>
<th>Explanatory Domains</th>
<th>Barthel Index</th>
<th>SF-36 PFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orpington Prognostic Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balance</td>
<td>51</td>
<td>34</td>
</tr>
<tr>
<td>Motor deficit in arm</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Cognition</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Proprioception</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>NIH Stroke Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arm strength</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>Leg strength</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Consciousness</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Sensory</td>
<td>2</td>
<td>...</td>
</tr>
<tr>
<td>Vision</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>Language</td>
<td>...</td>
<td>1</td>
</tr>
</tbody>
</table>

NIH indicates National Institutes of Health; PFI, physical functioning index. 
Values denote partial $R^2$ from linear regression analysis with a forward selection procedure with $P=0.05$ for entry criterion.
Prediction of Outcome After Stroke

training and certification for administration. The NIH Stroke Scale requires scoring of a greater number of aspects of neurological function, and it takes more than 10 minutes to complete the assessment. Also, because of the complexity of the NIH Stroke Scale, it is more likely to have missing items. Because of its additional assessment of a greater number of aspects of neurological function, in many cases such as aphasic patients the NIH Stroke Scale was not able to be used in 5 patients, resulting in incomplete evaluation of the patients. Conversely, all data items were ascertained for all stroke patients when the Orpington Prognostic Scale was used.

One limitation of our study is that the majority of our stroke subjects have mild to moderate stroke. By design, our study cohort did not have equal representation of severe strokes. Only 7% of our patients had an NIH Stroke Scale score of greater than 13, and 12% of the same cohort was categorized as major stroke by the Orpington Prognostic Scale. Therefore, the generalizability of the predictive value of the Orpington Prognostic Scale and the NIH Stroke Scale in this study may be limited to mild and moderate strokes.

In summary, clinicians and researchers who want to predict functional outcomes should select measures that are simple and do not require intense resources or training. Our results demonstrate that in a sample of mostly mild and moderate strokes, the Orpington Prognostic Scale compared with the NIH Stroke Scale is simpler to use and is a slightly better predictor of ADL and higher levels of physical function.

Acknowledgments

This study was funded by the Department of Veterans Affairs Rehabilitative Research and Development (E879RC) and Glaxo-Wellcome Pharmaceuticals. Participating facilities in the greater Kansas City area include the following: Baptist Hospital, Department of Veterans Affairs Medical Centers at Kansas City and Leavenworth, Liberty Hospital, Medical Center of Independence, Mid-American Rehabilitation Hospital, Rehabilitation Institute, Research Medical Center, St. Luke’s Hospital, St. Joseph Health Center, Trinity Lutheran Hospital, and University of Kansas Medical Center.

References


Prediction of Functional Outcome After Stroke: Comparison of the Orpington Prognostic Scale and the NIH Stroke Scale
Sue-Min Lai, Pamela W. Duncan and John Keighley

Stroke. 1998;29:1838-1842
doi: 10.1161/01.STR.29.9.1838

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/29/9/1838

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/