Outcome Measures in Acute Stroke Trials
A Systematic Review and Some Recommendations to Improve Practice
Pamela W. Duncan, PhD, FAPTA; Henrik Stig Jorgensen, MD, DMSci; Derick T. Wade, MD

Background—There is little consistency in the measurement of outcome in acute stroke trials, and this may complicate interpretation of the results and reduce the likelihood of detecting worthwhile drug effects. This study aims to investigate empirically the measures used to date and to give recommendations for future studies.

Summary of Comment—A systematic review of all published randomized studies of acute stroke drug intervention was undertaken, and the measures used were recorded. Fifty-one studies involving 57,214 subjects were identified. These studies used 14 different measures of impairment, 11 different measures of activity, 1 measure of “quality of life,” and 8 miscellaneous other measures. Timing of outcome assessments varied from 1 week to 1 year, with the modal time being 3 months. Many studies used ordinal measures but dichotomized results for analysis. Of the 51 studies included in the review, only 21 demonstrated benefit with the defined primary outcome measure. In several studies, however, post hoc analysis using varied outcome measures or varied cut points for dichotomizing outcomes resulted in positive results, whereas the primary study analysis failed to do so.

Conclusions—There is no consensus on the level of outcome to be used, the method of measurement to be used, or the most appropriate timing of the assessment. It is recommended that future studies should include extended/instrumental activities and advanced mobility as components of the primary outcome measure, with outcome assessment being undertaken at 6 months. New initiatives in developing stroke-specific outcomes may address some of the current problems in the assessment of stroke outcomes (Stroke. 2000;31:1429-1438.)

Key Words: outcome ■ stroke

Several recently reported studies of acute treatments for stroke have generated controversy because there were inconsistencies between various outcomes within each study.1–4 For example, the results of the European Cooperative Acute Stroke Study (ECASS) I study of recombinant tissue plasminogen activator (rtPA)5 were not positive; however, post hoc analysis of the same data using the outcome measures that were used in the National Institute of Neurological Disorders and Stroke (NINDS) rtPA trial produced favorable findings.2 The primary analysis in the subsequent ECASS II assessment of rtPA did not confirm a benefit, yet post hoc analysis that varied the cutoff points for favorable outcome on the modified Rankin scale did produce statistically significant results.3 Similarly, in the recent PROlyse in Acute Cerebral Thromboembolism (PROACT) II study of intra-arterial recombinant prourokinase,4 conflicting results were observed when the outcome measures varied.

The use of less than optimal methods for defining and measuring outcome may have caused these problems. Setting this in context, in the last 20 years there have been 51 phase II and phase III randomized clinical trials evaluating the efficacy of pharmacological interventions aimed at improving outcomes in individuals with stroke.1,3–52 Twenty-one of these trials have demonstrated benefit, but none has been confirmed or has significantly altered clinical management. The design of the trials and selection of outcome measures may have made detection of actual benefit less likely and the results controversial.

When considering the measurement of outcome, it is helpful to work within a framework or model of illness to categorize and classify the data that might be collected. The most widely used, and now generally accepted, model of illness is that promulgated by the World Health Organization (WHO), the International Classification of Impairments, Disabilities, and Handicaps (ICIDH).53 This system has recently been revised to provide health outcomes in terms of body, person, or social function and to provide a common framework for research.54 The new revision has changed terminology. The terms disability and handicap have been replaced by more positive terms, limitations in activities and restriction in participation. The revision also includes important contextual factors (Table 1).
Within the WHO ICIDH framework, outcomes may be measured at different levels. Any intervention will be expected to effect some specific change; in the case of most drugs, this is a reduction in the volume of brain damage, an effect at the level of pathology. However, the patient is more likely to value his or her ability to undertake a range of activities or to participate in social roles. Naturally, when selecting a measure, one must consider characteristics such as validity, reliability, and sensitivity to any differences expected. In addition, the items in any outcome measure used should relate to only a single level within the ICIDH because of the difficulty in interpreting results from a measure that reflects more than 1 level.

Assessment of stroke outcomes poses additional problems, given the natural history of the illness. Change may continue over months, and the rate and extent of change may vary between the different levels of the ICIDH. The final status of any particular patient cannot always be predicted with certainty, but prognostic factors relating to groups of patients are known for most outcomes. Specifying a favorable outcome is difficult, because most outcomes form a continuum. Consequently, studies must consider the timing of any assessments, the measurement of factors known to impact on the chosen outcomes, the likely statistical distribution of outcomes at the chosen end point, and the methods of statistical analysis best suited to the chosen outcome.

Consensus on and consistency in the use of appropriate measures of outcome and factors relating to outcome, the timing of measurements, and the methods of analysis would have many benefits for the medical community. The interpretation of results might be simplified. Meta-analysis would be much easier and more powerful. Comparison between studies would be facilitated. New studies could use published data to estimate the number of subjects needed. Readers of research would gain familiarity with the outcomes measured.

This article reports on an empirical analysis of published studies, investigating various aspects of outcome measurement in light of the above discussion. Specifically, the study (1) evaluates the use and timing of outcome measures in published drug trials, (2) examines the extent to which definitions of favorable and unfavorable outcomes and the timing of assessments are complicated by the natural history of stroke recovery, and (3) makes recommendations on how to choose measures for use in trials of drug treatment in the acute phase of stroke.

**Methods**

To identify stroke trials to be included in this synthesis, we undertook a systematic literature review of all phase II and phase III randomized trials of pharmacological interventions in the acute phase of stroke, published in English in 1980 or later. Multiple overlapping literature search strategies were used to identify a complete listing of trials. The MEDLINE database was searched by using combinations of the key words “cerebrovascular disease,” “stroke,” “trials,” and “outcomes.” We searched the Cochrane Collaboration’s comprehensive register of trials by using the key words “cerebrovascular disease.” In addition to the automated search strategies, reference lists and bibliographies of related journal articles and books were searched manually for additional trials. Care was taken to include each study only once when multiple publications referred to a single study.

From each of the identified studies, we abstracted the following data: intervention tested, number of subjects, subject inclusion and exclusion criteria, end points measured, timing of outcome measurements, defined cutoff points for outcomes, percentage of treatment group that achieved each defined outcome, percentage of placebo group that achieved each outcome, whether results were adjusted for subject severity, and the conclusion (results with P<0.05 were considered significant).

After data collection, each outcome measure was classified into one of the following categories: death or, at the level of pathophysiological parameters (blood pressure, laboratory values, and recanalization), impairment, activity, or participation. Measures were classified according to the system used by Roberts and Counsell, which includes the Rankin/modified Rankin scale as a measure of activity rather than participation.

Measurement instruments were also categorized as those with evidence on validity and reliability and those for which we could find no such evidence; it was specifically noted whether each measure had been validated for use with acute-stroke patients. To determine this, we reviewed references cited in the trials and published works that review the properties of outcome measures used in stroke research, and we undertook searches of the literature when the validity and reliability of measures could not be established from either of the first 2 methods.

<table>
<thead>
<tr>
<th>TABLE 1. Stroke Outcomes: WHO ICIDH-2 Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ICCIDH-2 Framework</strong></td>
</tr>
<tr>
<td>Illness of person</td>
</tr>
<tr>
<td>Pathology</td>
</tr>
<tr>
<td>Impairment</td>
</tr>
<tr>
<td>Activity (was disability)</td>
</tr>
<tr>
<td>Participation (was handicap)</td>
</tr>
<tr>
<td>Contextual factors</td>
</tr>
<tr>
<td>Person experiences</td>
</tr>
<tr>
<td>Physical environment</td>
</tr>
<tr>
<td>Social environment</td>
</tr>
<tr>
<td><strong>Synonym</strong></td>
</tr>
<tr>
<td>Disease/diagnosis</td>
</tr>
<tr>
<td>Symptoms/signs</td>
</tr>
<tr>
<td>Function/observed behavior</td>
</tr>
<tr>
<td>Social positions/roles</td>
</tr>
<tr>
<td>Prior function, comorbid disease</td>
</tr>
<tr>
<td>Home</td>
</tr>
<tr>
<td>Social support from family and friends</td>
</tr>
<tr>
<td><strong>Level of Description</strong></td>
</tr>
<tr>
<td>Organ/organ system</td>
</tr>
<tr>
<td>Body</td>
</tr>
<tr>
<td>Interaction of person and environment</td>
</tr>
<tr>
<td>Person in their social context</td>
</tr>
<tr>
<td><strong>Examples</strong></td>
</tr>
<tr>
<td>May affect stroke outcomes</td>
</tr>
<tr>
<td>May affect need for equipment and modifications</td>
</tr>
<tr>
<td>May affect motivation and stroke outcomes</td>
</tr>
<tr>
<td><strong>Comment</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Our search resulted in the identification of 51 trials, enrolling a total of 57,214 subjects. Study size ranged from 21 to 21,106 subjects, with a median enrollment of 287. Table 2 includes a description of study populations.

Inclusion and exclusion criteria for study subjects were variable, with much heterogeneity in age ranges included, severity of the current stroke, comorbid diseases, and pre-existing disability. The entry criteria for stroke trials may need to be variable because therapies are being tested for different populations and for different interventions. However, most of the studies included in the present review did not attempt to adjust the measured outcomes for age, stroke severity, or comorbidities.

The present review found considerable variability in the selection of outcome measures across studies. Of the 51 studies included in the present review, only 29 specifically defined the measure(s) and time frame(s) to be considered as their primary end points. When looking at all outcome measurements (primary or secondary), we observed variation in the levels of function selected for outcome measurement: ≈80% of the studies reviewed included a measure of impairment, ≈75% included a measure of activity, and only 1 study reported a measure of quality of life or health status. None of the studies reviewed included a measure of participation. Approximately two thirds of the studies reported death as an end point (Table 3).
Within functional levels of outcome selected for measurement, there was little consensus on the instruments used for measurement (Table 4). Of the 42 studies that measured impairment, 8 different stroke scale instruments, which reflect various neurological symptoms, were used. The modified Mathew scale was the most commonly used instrument; it was selected for use in 10 studies. Only 2 studies measured cognitive deficit, with both using the Mini-Mental State Examination. One study used a comprehensive neuropsychological battery, which included existing measures of cognitive and communication processes. Depression, a common sequela after stroke that has been shown to relate to recovery, was assessed in 3 trials.

There appears to be better consensus in the selection of measures of activity limitation, with the Barthel Index and the Rankin or modified Rankin scale being selected most frequently. Nine other activity measurement instruments were selected for use in at least 1 of the studies. In 8 of the 51 studies reviewed, investigators apparently did not find existing measurement instruments adequate, and new self-developed measures were used.

The majority of the instruments selected for outcome measurement have been shown to be reliable and valid for use with acute-stroke patients (Table 4). However, the impairment measure that was used most frequently, the modified Mathew scale, has been harshly criticized by numerous researchers for lack of validity, low interobserver reliability, and low internal consistency. The most commonly selected measure of activity, the Barthel Index, has been extensively tested and is considered to have good reliability and validity but is also known to be insensitive to small changes in functional status and to have significant ceiling effects. Another frequently used instrument, the Rankin handicap scale, is widely recognized as simple and reliable but has been criticized as inherently insensitive and for mixing objective and subjective items, which span impairment, activity, and participation aspects of recovery.

We found little agreement in definition of favorable or unfavorable outcomes. Although the Barthel Index was the most common measure, used in 27 of the studies that we reviewed, the cutoff scores used to differentiate favorable and unfavorable outcomes were defined in 7 different ways across the studies that used the measure. Similar variability was found across other measures used in the reviewed studies. Table 5 illustrates the cutoff points used and the large variability in the proportion of the placebo group that achieved the defined end point in each trial.

We also found tremendous variability in the selection of time frames for measurement of outcomes (see Figures 1 and 2). Although 3 months from baseline was the time frame most frequently selected for the primary outcome and for final follow-up measurement, the time points used varied from 1 day to 1 year for primary end points and from 1 week to 1 year for final measurements.

**Discussion**

The present review has emphasized that there is no consistency in the selection of outcome measures or the timing of assessments nor is there agreement on the definition or measurement of recovery. There is tremendous heterogeneity of patients enrolled in trials without adequate adjustment for expected outcomes. This systematic review has included all the major trials and is unlikely to have missed many significant studies. It highlights the need for some agreement on assessing stroke outcomes.

In the present state of outcome measurement in acute stroke trials, there are many unmet needs that occur because researchers have not incorporated a systematic framework for assessing outcomes. In addition, some of the existing outcome measures were not developed specifically for stroke survivors or with full understanding of the range of stroke recovery.

The WHO ICIDH-2 provides a useful framework. Within this framework, measures of impairment are the closest practical measures relating to the volume of brain loss, the pathophysiology that drugs presumably will reduce directly. Furthermore, impairment measures are probably the best markers of prognosis. Therefore, any study should include impairment level measures to assess case mix and as surrogate measures of outcome and should check to ensure that the expected neurological recovery has occurred. Impairment measures should not, however, be the primary outcome, because patients are more concerned with activity.

Measures at the level of activities (previously disability) are the most important primary outcome measure. Within this domain, measures of basic activities of daily living (ADL; eg, the Barthel ADL index) have significant weaknesses, as illustrated by the conflicting finding in the recent PROACT II study. In an unselected stroke population, ≈60% of the patients will make a “complete recovery,” scoring the top score, whereas a further 20% will make no recovery, remaining severely dependent. Thus, only ≈20% of patients have the potential to show a difference between groups in outcome, significantly reducing the power of any study. Consequently, researchers must consider recovery beyond ADL, especially if they are enrolling individuals with mild and moderate stroke.

Measures that assess a higher level of activity (ie, instrumental or extended ADL) or mobility may be more sensitive to differences between groups. There are several advantages to measures of ADL and mobility. They can be collected
equally reliably from the patient or from others; thus, there is little risk of missing data.\textsuperscript{71,72} They are relatively objective. They are simple and relevant to the patient, again increasing ease of full data collection. The main disadvantage is that the link between the extent of loss at the level of pathology and the level of activity is weak; therefore, other factors may influence the outcome.\textsuperscript{73,74} However, many of these factors (eg, diabetes, hyperthermia, depression, and social support)\textsuperscript{75} are known and can be allowed for in measures of case mix. Mobility measures, such as the Rivermead Mobility Index,\textsuperscript{62,76} capture a range of performance that is meaningful to the patient. Assessments of ranges of mobility skills (sitting and standing balance and assisted ambulation to advanced mobility, such as climbing stairs or ambulation in the community) do not suffer from ceiling effects.\textsuperscript{62,76}

Measurement of participation and quality of life both appear attractive but may pose problems. Measures of participation are only now being developed for stroke survivors.\textsuperscript{69,77} There is great debate about the construct of “quality of life,” and many of the measures said to assess the quality of life are also measuring activities or emotion.\textsuperscript{78} Furthermore, there are many extraneous factors that influence both participation and quality of life, making it less easy to detect the effect of the drug.\textsuperscript{74}

Measurement of emotion may be important, given the frequency of depression and its relation to activities, partici-
TABLE 5. Performance of Placebo Groups on Defined Categorical End Points

<table>
<thead>
<tr>
<th>Defined Cut Point</th>
<th>Study Reporting</th>
<th>% of Placebo Group Achieving</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favorable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥95</td>
<td>2, 10, 37, 46</td>
<td>23%, 38%, 33%, 35.4%</td>
</tr>
<tr>
<td>≥90</td>
<td>4, 38</td>
<td>35.7%, 32%</td>
</tr>
<tr>
<td>≥75</td>
<td>19, 25, 36</td>
<td>41%, 40.1%, 50%</td>
</tr>
<tr>
<td>≥60</td>
<td>4, 6, 12, 17, 43</td>
<td>58%, 70.1%, 49%, 54.8%, 47%</td>
</tr>
<tr>
<td>≥50</td>
<td>21</td>
<td>76%</td>
</tr>
<tr>
<td>60–90</td>
<td>43</td>
<td>19.4%</td>
</tr>
<tr>
<td>Unfavorable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60 or dead</td>
<td>20, 33</td>
<td>44.6%, 40%</td>
</tr>
</tbody>
</table>

Rankin

| Favorable         |                 |                              |
| Full recovery 0    | 14              | 37.7%                        |
| Full recovery ≤1   | 2, 3, 10, 38    | 16%, 28%, 36.6%, 21.4%       |
| Independent 1–2    | 14              | 33.6%                        |
| Independent ≤2     | 4, 19, 25       | 31%, 29.7%, 25%              |

Unfavorable

| 2               | 11, 19, 21, 33 | 39%, 48%, 65%, 53%          |
| 3 or death       | 13, 14         | 81.8%, 31.6%                |

NIHSS

| 1               | 2, 4, 10, 37, 38| 12%, 22%, 16%, 7%, 12%       |
| 7               | 25             | 47.3%                        |

GOS

| 1               | 10             | 19%                          |

Combined GOS and BI

| Very favorable   |                 |                              |
| GOS 1 and BI     | 16              | 47.0%                        |
| 19–20            |                 |                              |
| Favorable        |                 |                              |
| GOS 1 or 2 and   | 16              | 73.7%                        |
| BI ≥12           |                 |                              |

Modified Mathew

| Favorable        |                 |                              |
| ≥75              | 36              | 54%                          |

MMSE

| Favorable        |                 |                              |
| ≥25              | 37              | 34.5%                        |

BI indicates Barthel Index; GOS, Glasgow Outcome Scale; and MMSE, Mini-Mental State Examination.

from the perspective of the patient and caregiver, and they include domains that tap the full impairment/activity/participation continuum and the frequently omitted areas of language, cognition, and hand function. Both of these measures include a broader range of activities and mobility. Early evaluations of the psychometric properties of both instruments are promising, and the SIS demonstrates good sensitivity to change in patients with mild and moderate stroke. The SIS captures 8 domains that are individually scored. Four of the SIS domains (strength, hand function, ADL/instrumental ADL, and mobility) may be averaged to create a physical domain, but it is invalid to sum across the domains of communication, memory and thinking, emotion, and participation.

The time course of recovery from stroke must be considered when selecting the time of assessment. Furthermore, it is important to distinguish between recovery at the impairment, activity, and participation levels. The degree and timing of recovery are known to relate to the initial severity of the stroke. In the Copenhagen Stroke Study, patients were stratified according to initial stroke severity. Functional recovery (measured by the Barthel Index) was achieved within 2 months from stroke onset by 95% of the patients with mild strokes (41% of the population), within 3 months in patients with moderate strokes (26%), within 4 months in patients with severe strokes (14%), and within 5 months after stroke onset in patients with the most severe strokes (19%). Neurological recovery (as measured by the Scandinavian Stroke Scale) preceded functional recovery by a mean of 2 weeks.

Five to 6 months after stroke seems an appropriate time point at which to measure neurological and functional outcome. Spontaneous recovery does not plateau until 5 to 6 months after stroke, especially in more severe strokes. Therefore, unless time to recovery is an important outcome, the
primary outcomes should occur when spontaneous recovery has plateaued in the placebo group. Measurement of participation should preferably take place at a time when the patient’s social condition has stabilized. Outcome measurement at the time of discharge from completed treatment and rehabilitation should be avoided because it limits comparability across studies because of geographic variations in treatment practices and lengths of stay and may be too early in the recovery process to detect maximal recovery.

Cutoff points for defining favorable or unfavorable outcomes are often arbitrarily selected. A recent review by Sulter et al\(^{80}\) of the use of the Barthel Index and modified Rankin scale in stroke trials found substantial variability in cutoff-point definition across trials and asserted that many of the cutoff points had been arbitrarily established and not validated. Similarly, we have found a lack of consistency in the selection of cutoff points in the trials reviewed in the present study. This lack of consistency in definition of a clinically meaningful recovery hinders the interpretation of results and comparison across trials. Limited evidence does exist to guide the researcher in the appropriate selection of cutoff scores. Granger et al\(^{81}\) have suggested that a score of 60 on the Barthel Index corresponds to the shift from dependence to assisted independence. In the same study, a score of 85 was found to correspond to the transition from minimal assistance to independence. Other studies\(^{82,83}\) have suggested that a Barthel score <85 corresponds to the point at which patients reported requiring assistance in ADL.

For the modified Rankin scale, Kay et al\(^{82}\) compared self-reported dependence with Rankin scores and found that a cutoff score of \(\geq 2\) corresponded with 85% sensitivity and 87% specificity to dependence, whereas a score of \(\geq 3\) corresponded to dependency with 94% sensitivity and 70% specificity. Sulter et al\(^{80}\) suggest that poor outcome may be easier to define than favorable outcome, that a single scale may be inadequate to interpret outcome, and that an appropriate definition may be that poor outcome exists if any of the following occur: death, institutionalization due to stroke, modified Rankin score >3, or Barthel Index score <60. However, if the measurement focus is limited to unfavorable outcomes, one cannot look at improvements, and we may fail to detect the shifting of patients to higher levels of function. Dichotomizing scales to define favorable and unfavorable outcomes reduces outcome information and may limit our ability to detect a significant shift in disability. Additionally, varying the cutoff points and selecting different levels of measurement (impairment, activity, or participation) significantly shifts the percentage of individuals who are deemed to have a good outcome.\(^{84}\) Finally, we may not know the clinical relevance of the dichotomization. Dichotomizing outcomes should be avoided; instead, statistical analysis that uses the whole range of data collected should be used. Because most of the outcome data are ordinal and are not normally distributed, nonparametric statistics should be used.

The results of the present review of drug trials in acute stroke demonstrate that in the past 20 years there has been little consistency in the selection of outcome measurement methods or in the definition of favorable or unfavorable outcomes in acute stroke trials. Instruments selected for outcome measurement are frequently inadequate. Clearly, when there is so much variability in the methods used for outcome assessment of randomized drug trials for acute stroke, it is difficult to know what to make of the results. Comparisons across trials are difficult because of inconsistent measure selection, varying definition of favorable or unfavorable outcomes, varying times for measurement, and differences in the inclusion and exclusion criteria for study subjects. The result may be that outcome measurement methods are contributing to the difficulty in the demonstration of treatment benefits. The most recently completed trials of rtPA serve as examples of the confusion in selecting outcomes. To improve the selection of stroke outcome measures, the following recommendations should be considered: (1) Primary outcome measures should be at the level of activities, capturing not only basic ADL but also including instrumental activities of daily living and advanced mobility. (2) Impairment outcome measures should be included to assess whether the drug has affected neurological recovery, the supposed primary level of action. (3) Assessment of the individual’s emotion should be considered, but because many individuals may not be able to respond and because proxy assessment of emotion is not reliable, there may be substantial missing data, rendering emotion as a poor primary outcome. (4) All outcome measures should have established psychometric properties (reliable, valid, and sensitive to change) and should have been tested in individuals with stroke. (5) Definition of recovery should not be dichotomized but rather should assess shifts in disability by use of nonparametric statistics. (6) The natural history of recovery of the population under study must be considered when outcome measures are selected. (7) Primary outcomes should be assessed at 6 months, especially in samples that include severe stroke. (8) Data collection should include baseline characteristics that could confound outcomes (eg, diabetes, body temperature, glucose levels, depression, and social support). (9) New initiatives in developing stroke-specific outcomes,\(^{85,77}\) which have incorporated the WHO ICIDH model and considered the range of stroke-related impairments and activities for different levels of stroke severity, may address some of the current problems in the assessment of stroke outcomes.

Acknowledgments
This study was supported by GlaxoWellcome, Inc, and the University of Kansas Claude D. Pepper Older Americans Independence Center, which is funded by the National Institute of Aging (P60 AG-14635-02). We wish to thank Barbara LaClair, MS, of the Kansas City, Mo, Department of Veterans Affairs Medical Center, for her help in researching and preparing this manuscript.

References
2. Hacke W, Bluhmki E, Steiner T, Tatlisumak T, Mahagne MH, Sacchetti ML, Meier D. Dichotomized efficacy end points and global end-point
1436 Stroke June 2000

10. Tissue plasminogen activator for acute ischemic stroke: the National
11. Thrombolytic Therapy with Streptokinase in Acute Ischemic Stroke: the CAST: randomised placebo-controlled trial of early aspirin use in
17. Low molecular weight heparinoid, ORG 10172 (danaparoid), and
18. Low molecular weight heparinoid, ORG 10172 (danaparoid), and
Outcome Measures in Acute Stroke Trials

Duncan et al


100. Granger CV, Hamilton BB, Gresham GE, Kramer AA. The stroke rehabilitation outcome study, II: relative merits of the total Barthel index score and a four-item subscore in predicting patient outcomes. Arch Phys Med Rehabil. 1989;70:100–103.


Outcome Measures in Acute Stroke Trials: A Systematic Review and Some Recommendations to Improve Practice
Pamela W. Duncan, Henrik Stig Jørgensen and Derick T. Wade

Stroke. 2000;31:1429-1438
doi: 10.1161/01.STR.31.6.1429

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/31/6/1429