Use of the Barthel Index and Modified Rankin Scale in Acute Stroke Trials

Geert Sulter, MD; Christel Steen, MS; Jacques De Keyser, MD, PhD

Background and Purpose—The Barthel Index (BI) and the Modified Rankin Scale (MRS) are commonly used scales that measure disability or dependence in activities of daily living in stroke victims. The objective of this study was to investigate how these scales were used and interpreted in acute stroke trials.

Methods—We identified from MEDLINE the major efficacy trials with neuroprotective drugs, thrombolytic drugs, and anticoagulants in acute ischemic stroke published between January 1995 and December 1998. We selected those trials that used the BI and/or MRS as outcome parameters.

Results—Fifteen trials fulfilling the inclusion criteria were identified. The BI was used in 13 and the MRS in 8. In 4 trials mean and median scores of the BI were used, and in 1 trial median scores of the MRS were compared. Primary end points included the BI in 7, the MRS in 6, and both the BI and MRS in 3. With regard to the BI, a variety of sum scores between 50 and 95 were used as cutoff scores to define favorable outcome. Favorable outcome on the MRS was defined as either ≤1 or ≤2.

Conclusions—Among the efficacy trials in acute stroke, we found remarkable differences in the choice of primary end points and in the definition of favorable outcome on both the BI and MRS. This lack of consensus strongly hinders the design, interpretation, and comparison of acute stroke trials. In general, it may be easier to define poor outcome instead of favorable outcome. Poor outcome could be defined if any of the following end points are reached: death, institutionalization due to stroke, MRS >3, or BI <60. (Stroke. 1999;30:1538-1541.)

Key Words: disability evaluation ■ outcome ■ stroke trials

S ubstantial efforts are being made to develop drug treatments that may minimize brain damage and improve outcome in patients with ischemic stroke. In clinical trials, the Barthel Index (BI) and the Modified Rankin Scale (MRS) are commonly used scales to assess outcome.

The BI was developed in 19651 and later modified by Granger and coworkers2 as a scoring technique that measures the patient’s performance in 10 activities of daily life. The BI is considered a reliable disability scale for stroke patients.3 The items can be divided into a group that is related to self-care (feeding, grooming, bathing, dressing, bowel and bladder care, and toilet use) and a group related to mobility (ambulation, transfers, and stair climbing). The maximal score is 100 if 5-point increments are used, indicating that the patient is fully independent in physical functioning. The lowest score is 0, representing a totally dependent bedridden state. The MRS measures independence rather than performance of specific tasks.4 In this way, mental as well as physical adaptations to the neurological deficits are incorporated. The scale consists of 6 grades, from 0 to 5, with 0 corresponding to no symptoms and 5 corresponding to severe disability (Table 1).

Although both scales are easy to use and have an acceptable degree of reliability,3,4 there is apparently no consensus on how these scales should be used to determine outcome in clinical trials. The aim of our study was to investigate how both scales were used and interpreted in recently published clinical drug trials in acute ischemic stroke.

Methods

We searched through MEDLINE for the major pivotal randomized trials (phase III trials) in acute stroke in which the efficacy of either thrombolytics, neuroprotective drugs, or antithrombotic compounds was investigated and the final results were published between January 1995 and December 1998. We selected those trials that used the BI and/or MRS either as primary or secondary end point. We assessed how the scores obtained from these scales were presented and interpreted.

Results

We identified 15 trials that together enrolled 9082 patients (Table 2). There were a variety of primary end points, including stroke scales, mortality, BI, MRS, and the Glasgow Outcome Scale (GOS). Four trials defined >1 primary end point. Assessment of the primary end points occurred at 3 months in 10 studies. In 1 study patients were scored at 1 month; in 4 other trials they were scored at 6 months.

The BI was used in 13 trials and served as primary outcome measure in 7 trials. Two trials compared mean scores, and 2
Granger and coworkers2 found that a score of 60 was a
been conducted on the clinical relevance of the sum scores.
ously, although the BI is widely used, few studies have
arbitrarily chosen and have never been validated. Curi-
to
being patients with a favorable outcome varied substan-
tially from trial to trial; a variety of sum scores from
varied from 3 to 6 categories (Table 3).

The MRS was used in 8 trials and was a primary outcome
parameter in 6. In 1 trial median scores of the MRS were
compared between the 2 treatment groups. Favorable out-
come on the MRS was defined in 3 trials as a score ≤1 and
in 5 trials as a score ≤2. In the second European Coopera-
tive Acute Stroke Study (ECASS II) trial, the proportion of
patients with a 3-months MRS of ≤1 was used to assess
efficacy of recombinant tissue plasminogen activator (r-
tPA).15 There was no statistically significant difference with
placebo. However, by shifting MRS grades from ≤1 to ≤2
(not a predefined end point) to indicate favorable outcome, a
significant result in favor of r-tPA was obtained.

Discussion
Over the last 4 years, a number of pivotal acute stroke trials
with thrombolytics, anticoagulants, and neuroprotective
drugs were published. All failed to demonstrate a significant
clinical beneficial effect on predefined primary end points,
except for the National Institute of Neurological Disorders
and Stroke (NINDS) and Fraxiparine in Ischemic Stroke
Study (FISS) trials.12,16 Remarkably, the criteria by which
outcome was measured and the measuring instruments that
were used differed considerably between the trials. The lack
of consensus regarding what should be considered a clinically
meaningful effect on outcome in stroke patients strongly
hinders the design, interpretation, and comparison of acute
stroke trials.

The BI was the most commonly used scale for assessing
activities of daily living. However, the criteria for classi-
fying patients with a favorable outcome varied substan-
tially from trial to trial; a variety of sum scores from ≥50
to ≥95 were used. In fact, many of the cutoff scores were
arbitrarily chosen and have never been validated. Curi-
ously, although the BI is widely used, few studies have
been conducted on the clinical relevance of the sum scores.
Granger and coworkers2 found that a score of 60 was a
pivotal score at which patients move from assisted inde-
pendence to dependence. In practical terms, with a score of
≥60, most patients were independent for essential personal
care, such as moving around unassisted, sphincter control,
eating, and personal toilet. A score of 85 usually corre-
sponded to independence with minimal assistance.2 This
means that the majority of patients were able to get dressed
and to move from armchair to bed unassisted. Kay and
coworkers20 and Dennis and colleagues21 also found that a
score of <85 corresponded to a state in which patients
reported needing help in performing activities of daily
living, with a sensitivity of 94% to 95% and a specificity
of 80% to 86%. Therefore, it is difficult to defend why in
the NINDS and ECASS II trials a BI of ≥95 (defined as
minimal or no disability) was chosen to define favorable
outcome.15,16 Why not use a score of ≥85, which has been
shown to correspond to an acceptable level of autonomy in
the majority of patients, or a score of 100 if one predicts
that the therapy should result in complete recovery? Some
investigators have even invented a combined BI/Rankin
Scale with a sum score of 110, although such a scoring
system has also not been functionally justified.6,14,15 The
BI is an ordinal (noncontinuous) scale. Therefore, para-
metric statistical methods cannot be used. We found it
cumbersonsome that prestigious journals accepted the presen-
tation of results in mean or median BI values,6,14,15,17
which are completely inappropriate statistical end points.
A number of studies distinguished between different cat-
egories of disability on the BI.6,8,9,11,16 This may allow
assessment of a global shift toward independence in
subsequent categories of scores rather than reliance on a
single score that dichotomizes outcome into favorable
or poor. However, the cutoff scores used to define these
different categories should also be validated and not
arbitrarily chosen.

In contrast to the BI, we found fewer variations between
the trials in use of the MRS. Only 2 cutoff scores were used:
≤1 in 3 trials investigating r-tPA and ≤2 in 5 other trials.
A striking example of how data can be differently interpreted by
post hoc analyses was recently demonstrated in the ECASS II
trial.15 It was possible, simply by shifting the cutoff score
from the predefined ≤1 grades to ≤2 grades on the MRS, to
obtain a statistically significant effect in favor of r-tPA. At
first glance, a cutoff score of ≤2 (able to look after own
affairs without assistance) appears to be more meaningful
than a score of ≤1 in terms of independence (Table 1). Kay
et al20 found that self-reported dependence had a sensitivity
of 85% and specificity of 87% against the MRS dichotomized at
≥2 and a sensitivity of 94% and specificity of 70% against
the MRS dichotomized at ≥3.

Concordance between outcome parameters is also impor-
tant in the evaluation of trial results. Although cutoff scores
were chosen on a pragmatic basis, this was one of the
strengths of the NINDS trial, in which each of the 4 outcome
parameters (BI, MRS, National Institutes of Health Stroke
Scale, GOS) showed consistency in the degree of benefit with
r-tPA.16 Some investigators advocate the use of self-reported
dependence as a valid means of dichotomizing stroke patients
for the purpose of clinical trials.20,21 We should be careful
with this definition of dependence. Many factors unrelated to

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No symptoms</td>
</tr>
<tr>
<td>1</td>
<td>No significant disability, despite symptoms; able to perform all usual duties and activities</td>
</tr>
<tr>
<td>2</td>
<td>Slight disability; unable to perform all previous activities but able to look after own affairs without assistance</td>
</tr>
<tr>
<td>3</td>
<td>Moderate disability; requires some help, but able to walk without assistance</td>
</tr>
<tr>
<td>4</td>
<td>Moderately severe disability; unable to walk without assistance and unable to attend to own bodily needs without assistance</td>
</tr>
<tr>
<td>5</td>
<td>Severe disability; bedridden, incontinent, and requires constant nursing care and attention</td>
</tr>
</tbody>
</table>

TABLE 1. Modified Rankin Scale

© 2017 American Heart Association
the intervention, such as psychological and socioeconomic factors, may influence this subjective statement. For example, a wheelchair-bound patient may have a level of income or a health insurance system that allows the acquisition of expensive technical aids to compensate for restrictions to perform desired activities, thereby making the patient “independent.” The BI and MRS are reliable measures that provide a more objective assessment of functional recovery after stroke.

TABLE 2. Pivotal Trials in Acute Ischemic Stroke Published in Final Form Between January 1995 and December 1998

<table>
<thead>
<tr>
<th>Clinical Trial (Year of Publication)</th>
<th>Compound Tested</th>
<th>No. of Randomized Patients</th>
<th>Primary End Point(s)</th>
<th>Time Interval Between Stroke Onset and Assessments</th>
<th>Favorable Outcome on BI</th>
<th>Favorable Outcome on MRS</th>
<th>Results of Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIST5 (1996)</td>
<td>Flunarizine</td>
<td>331</td>
<td>Neurological status, BI, MRS</td>
<td>6 mo ≥50 (independence)</td>
<td>≥2</td>
<td>No improvement of neurological and functional outcome</td>
<td></td>
</tr>
<tr>
<td>PASS6 (1997)</td>
<td>Piracetam</td>
<td>927</td>
<td>Neurological status</td>
<td>1 mo Comparison of mean scores +110 (complete recovery)</td>
<td>...</td>
<td>No effect on neurological outcome</td>
<td></td>
</tr>
<tr>
<td>RANTTAS7 (1996)</td>
<td>Tirilazad</td>
<td>658</td>
<td>BI, GOS</td>
<td>3 mo ≥60</td>
<td>...</td>
<td>No improvement of functional outcome</td>
<td></td>
</tr>
<tr>
<td>LUB-INT-98 (1997)</td>
<td>Lubeluzole</td>
<td>721</td>
<td>Mortality</td>
<td>3 mo ≥75</td>
<td>≤2</td>
<td>No significant reduction in mortality but improvement of functional outcome</td>
<td></td>
</tr>
<tr>
<td>LUB-INT-59 (1998)</td>
<td>Lubeluzole</td>
<td>725</td>
<td>Mortality</td>
<td>3 mo ≥75</td>
<td>≤2</td>
<td>No effect on mortality or functional outcome</td>
<td></td>
</tr>
<tr>
<td>ASSIST10 (1997)</td>
<td>Selfotel (CGS 19755)</td>
<td>509</td>
<td>BI</td>
<td>90 d ≥60</td>
<td>...</td>
<td>Trials stopped prematurely because of unfavorable risk-benefit ratio</td>
<td></td>
</tr>
<tr>
<td>Ebselen in Acute Stroke11 (1998)</td>
<td>Ebselen</td>
<td>302</td>
<td>GOS</td>
<td>1 and 3 mo ≥75 (mild disability)</td>
<td>...</td>
<td>No effect at 3 mo, but significantly more good recovery or moderate disability at 1 mo</td>
<td></td>
</tr>
<tr>
<td>FISS12 (1995)</td>
<td>Low-molecular-weight heparin</td>
<td>312</td>
<td>Self-reported dependence</td>
<td>6 mo ≥85 (independence)</td>
<td>...</td>
<td>Significantly less death+dependence at 6 mo; no effect at 3 mo</td>
<td></td>
</tr>
<tr>
<td>TOAST13 (1998)</td>
<td>Low-molecular-weight heparinoid</td>
<td>1281</td>
<td>GOS, BI</td>
<td>3 mo ≥60</td>
<td>...</td>
<td>No improvement of functional outcome</td>
<td></td>
</tr>
<tr>
<td>ECASS14 (1995)</td>
<td>r-tPA</td>
<td>620</td>
<td>BI, MRS</td>
<td>3 mo Comparison of median scores; use of a combined BI/Rankin Scale (110 points)</td>
<td>Comparison of median scores +1</td>
<td>...</td>
<td>No improvement of functional outcome; may be beneficial in a well-defined subgroup</td>
</tr>
<tr>
<td>ECASS II15 (1998)</td>
<td>r-tPA</td>
<td>800</td>
<td>MRS</td>
<td>3 mo Comparison of median scores including use of a combined BI/Rankin Scale (110 points) + ≥95</td>
<td>≤1 (post hoc also ≥2)</td>
<td>...</td>
<td>No statistical benefit on predefined outcome</td>
</tr>
<tr>
<td>NINDS16 (1995) (part 1)</td>
<td>r-tPA</td>
<td>291 + 333 (part 2)</td>
<td>BI, MRS, GOS, NIHSS</td>
<td>3 mo ≥95</td>
<td>≤1</td>
<td>r-tPA improves clinical outcome at 3 mo</td>
<td></td>
</tr>
<tr>
<td>MAST-E17 (1996)</td>
<td>Streptokinase</td>
<td>310</td>
<td>Death+severe disability on MRS</td>
<td>6 mo Comparison of mean scores</td>
<td>≤2</td>
<td>Increase in mortality</td>
<td></td>
</tr>
<tr>
<td>MAST-I18 (1995)</td>
<td>Streptokinase</td>
<td>622</td>
<td>MRS</td>
<td>6 mo ...</td>
<td>≤2</td>
<td>Increase in 10-d mortality; no significant beneficial effect at 6 mo</td>
<td></td>
</tr>
<tr>
<td>ASK19 (1996)</td>
<td>Streptokinase</td>
<td>340</td>
<td>Mortality + BI</td>
<td>3 mo ≥60</td>
<td>...</td>
<td>Increased mortality and morbidity</td>
<td></td>
</tr>
</tbody>
</table>

Another important issue is that in the analysis of the trials, different types of stroke were mixed together. Bamford and coworkers showed that the 6-month combined outcome of being dead or dependent (defined as MRS score >2) differed greatly between patients with total anterior circulation infarcts and those with lacunar or partial anterior circulation infarcts. Efficacy in a controlled trial should cover the entire study population. However, although effective randomization should ensure a good balance of stroke subtypes in the treated and placebo groups, it would also be appropriate to analyze outcome separately in function of stroke subtypes. If patients with anterior circulation lesions are selected, one should distinguish between total anterior circulation infarction, partial anterior circulation infarction, and lacunar infarction.

Instead of trying to define favorable outcome, for which there will be no consensus, we believe that it may be easier to define poor outcome. Rather than using a single scale, we suggest a definition of poor outcome for each of the following occurs: death, institutionalization due to stroke, MRS 3, or BI <60. Whereas there is no consensus regarding the definition of favorable outcome, there will be less disagreement that each of these 4 conditions corresponds to an unfavorable outcome, and it is a more objective instrument than self-reported dependence.

Acknowledgment
This study was supported by a grant from the Academic Hospital Groningen (to Dr Sulter).

References
Use of the Barthel Index and Modified Rankin Scale in Acute Stroke Trials
Geert Sulter, Christel Steen and Jacques De Keyser

*Stroke*. 1999;30:1538-1541
doi: 10.1161/01.STR.30.8.1538

*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1999 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/30/8/1538

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/