Optimizing Cutoff Scores for the Barthel Index and the Modified Rankin Scale for Defining Outcome in Acute Stroke Trials

Maarten Uyttenboogaart, MD; Roy E. Stewart, MSc; Patrick C.A.J. Vroomen, MD, PhD; Jacques De Keyser, MD, PhD; Gert-Jan Luijckx, MD, PhD

Background and Purpose—There is little agreement on how to assess outcome in acute stroke trials. Cutoff scores for the Barthel Index (BI) and modified Rankin Scale (mRS) are frequently arbitrarily chosen to dichotomize favorable and unfavorable outcome. We investigated sensitivity and specificity of BI cutoff scores in relation to the mRS to obtain the optimal corresponding BI and mRS scores.

Methods—BI and mRS scores were collected from 1034 ischemic stroke patients. Sensitivity and specificity were calculated for BI cutoff scores from 45 to 100 in mRS score 1, 2, and 3 and were plotted in receiver operator characteristic (ROC) curves.

Results—The cutoff scores for the BI with the highest sum of sensitivity and specificity were 95 (sensitivity 85.6%; specificity 91.7%), 90 (sensitivity 90.7%; specificity 88.1%), and 75 (sensitivity 95.7%; specificity, 88.5%) for, respectively, mRS 1, 2, and 3. The area under the ROC curve was 0.933 in mRS 1, 0.960 in mRS 2, and 0.979 in mRS 3.

Conclusions—The optimal cutoff scores for the BI were 95 for mRS 1, 90 for mRS 2, and 75 for mRS 3. For future acute stroke trials that assess stroke outcome with the BI and mRS, we recommend the use of these BI cutoff score(s) with the corresponding mRS cutoff score(s), to ensure the use of consistent and uniform end points. (Stroke. 2005;36: 1984-1987.)

Key Words: disability evaluation ■ outcome assessment ■ stroke

Several randomized controlled acute stroke trials have been designed to investigate effectiveness of therapeutic interventions. A major point of discussion is how to define outcome in acute stroke trials with disability and handicap scales.1–6 The most widely used scales are the modified Rankin Scale (mRS) and the Barthel Index (BI).

The mRS has proved to be valid and reliable for defining outcome in stroke patients.7,8 Although the mRS was designed as a handicap scale,9 it should be considered a disability scale.10 The mRS defines 6 different grades of disability, from 0 for “no symptoms at all” to 5 for “severe disability or bedridden, incontinent, and requiring constant nursing care and attention,” and grade 6 for death.

The BI has also shown to be valid and reliable for assessing disability in stroke patients.8,11 It contains 10 items with varying weights that score activities of daily living (ADL). The items bathing and grooming are scored 0 or 5; the items feeding, dressing, controlling bladder, controlling bowel, getting onto and off the toilet, and ascending and descending stairs are scored 0, 5, or 10. Items regarding moving from wheelchair to bed and walking on level surface are scored 0, 5, 10, or 15. The total BI is a cumulative score of the 10 items, with a maximum score of 100 corresponding with complete independence, and a minimum score of 0 corresponding with total dependence.

There is little consensus on the optimal implementation of the BI and mRS as outcome measure in acute stroke trials. It is unclear which outcome scale is preferable. Moreover, the cutoff scores distinguishing favorable and unfavorable outcome are highly variable between various acute stroke trials.4 This issue has great consequences for the design and interpretation of acute stroke trials. The BI has a larger score range and therefore more possible cutoff scores compared with the mRS. Less is known which BI scores are corresponding with the different mRS scores. There have been only a few studies that determined pivotal BI cutoff scores, and none of them were related to the mRS.12–14 In this article, we investigated which cutoff scores on the BI corresponded to mRS grades 1, 2, and 3.
Results

Population Characteristics
From the 1034 patients, 547 (52.9%) were female. The mean age was 69.1 years (SD 12.8 years). Median BI score was 80 with an interquartile range from 40 to 100. The mRS score distribution was mRS 0, 9.1%; mRS 1, 17.8%; mRS 2, 13.1%; mRS 3, 19.1%; mRS 4, 29.7%; and mRS 5, 11.2%.

Sensitivity and Specificity of BI and mRS Cutoff Scores
The sensitivity and specificity for the cutoff scores of the BI in relation to mRS 1, 2, and 3 were calculated (Table 2) and plotted in ROC curves (Figure).

For mRS 1, the optimal cutoff score on the BI was 95, with a sensitivity of 85.6% (95% CI, 82.9% to 87.9%) and a specificity of 91.7% (95% CI, 87.8% to 94.5%). For mRS 2, the BI score with the highest sum of sensitivity and specificity was 90, with a sensitivity of 90.7% (95% CI, 88.1% to 92.7%) and a specificity of 88.1% (95% CI, 84.6% to 90.9%). An mRS 3 agreed most with a BI score of 75, with a sensitivity of 95.7% (95% CI, 93.3% to 97.5%) and a specificity of 88.5% (95% CI, 85.8% to 90.8%). In all 3 mRS cutoff scores, sensitivity (rate true unfavorable outcome) increased and specificity (rate true favorable outcome) decreased when BI scores increased.

Subsequently, AUCs were calculated. The AUC for the BI cutoff scores was 0.932 (95% CI, 0.916 to 0.949) in mRS 1, 0.960 (95% CI, 0.949 to 0.971) in mRS 2, and 0.979 (95% CI, 0.972 to 0.985) in mRS 3.

Discussion
In this study, we analyzed the optimal cutoff scores for the BI and the mRS. These were found to be BI 95 for mRS 1, BI 90 for mRS 2, and BI 75 for mRS 3. This finding may have

### TABLE 1. Calculation of Sensitivity and Specificity for BI Cutoff Scores in mRS 1, 2, and 3

<table>
<thead>
<tr>
<th>mRS Reflecting Favorable Outcome</th>
<th>mRS Reflecting Unfavorable Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI reflecting favorable outcome (BI ≥ 45–100)</td>
<td>A</td>
</tr>
<tr>
<td>BI reflecting unfavorable outcome (BI &lt; 45–100)</td>
<td>C</td>
</tr>
</tbody>
</table>

A, indicates true favorable outcome; B, false favorable outcome; C, false unfavorable outcome; D, true unfavorable outcome.

Sensitivity, D/(D + B); specificity, A/(A + C).

### TABLE 2. Sensitivity, Specificity, and Sum Score for BI Cutoff Scores in mRS 1, 2, and 3

<table>
<thead>
<tr>
<th>BI score</th>
<th>100</th>
<th>95</th>
<th>90</th>
<th>85</th>
<th>80</th>
<th>75</th>
<th>70</th>
<th>65</th>
<th>60</th>
<th>55</th>
<th>50</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRS 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.927</td>
<td>0.856</td>
<td>0.786</td>
<td>0.706</td>
<td>0.660</td>
<td>0.620</td>
<td>0.573</td>
<td>0.529</td>
<td>0.480</td>
<td>0.433</td>
<td>0.390</td>
<td>0.349</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.813</td>
<td>0.917</td>
<td>0.935</td>
<td>0.960</td>
<td>0.971</td>
<td>0.978</td>
<td>0.993</td>
<td>0.993</td>
<td>0.996</td>
<td>0.996</td>
<td>0.996</td>
<td>0.996</td>
</tr>
<tr>
<td>Sum</td>
<td>1.740</td>
<td>1.773</td>
<td>1.721</td>
<td>1.666</td>
<td>1.631</td>
<td>1.598</td>
<td>1.566</td>
<td>1.522</td>
<td>1.476</td>
<td>1.429</td>
<td>1.386</td>
<td>1.345</td>
</tr>
<tr>
<td>mRS 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.981</td>
<td>0.952</td>
<td>0.907</td>
<td>0.837</td>
<td>0.794</td>
<td>0.750</td>
<td>0.696</td>
<td>0.643</td>
<td>0.585</td>
<td>0.527</td>
<td>0.475</td>
<td>0.425</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.651</td>
<td>0.809</td>
<td>0.881</td>
<td>0.939</td>
<td>0.966</td>
<td>0.978</td>
<td>0.993</td>
<td>0.993</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td>Sum</td>
<td>1.632</td>
<td>1.761</td>
<td>1.789</td>
<td>1.776</td>
<td>1.760</td>
<td>1.728</td>
<td>1.689</td>
<td>1.636</td>
<td>1.583</td>
<td>1.525</td>
<td>1.473</td>
<td>1.423</td>
</tr>
<tr>
<td>mRS 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.991</td>
<td>0.976</td>
<td>0.957</td>
<td>0.917</td>
<td>0.875</td>
<td>0.813</td>
<td>0.745</td>
<td>0.676</td>
<td>0.619</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.460</td>
<td>0.596</td>
<td>0.691</td>
<td>0.794</td>
<td>0.846</td>
<td>0.885</td>
<td>0.923</td>
<td>0.948</td>
<td>0.967</td>
<td>0.979</td>
<td>0.984</td>
<td>0.995</td>
</tr>
<tr>
<td>Sum</td>
<td>1.460</td>
<td>1.596</td>
<td>1.691</td>
<td>1.785</td>
<td>1.822</td>
<td>1.842</td>
<td>1.840</td>
<td>1.823</td>
<td>1.780</td>
<td>1.724</td>
<td>1.660</td>
<td>1.614</td>
</tr>
</tbody>
</table>

*Maximum sum of sensitivity and specificity in mRS 1; †maximum sum of sensitivity and specificity in mRS 2; ‡maximum sum of sensitivity and specificity in mRS 3.
consequences for the definition of outcome in acute stroke trials.

A recent acute stroke trial defined favorable outcome with an mRS ≤2 and BI ≥75.\textsuperscript{18} According to our results, these cutoff scores could be suboptimal. The sensitivity (75.0%) and specificity (97.8%) of these cutoff scores implicates that 25% percent of the patients would have a favorable outcome according to the BI but an unfavorable outcome according to the mRS. With regard to the specificity, 2.2% would have an unfavorable outcome according to the BI but a favorable outcome according to the mRS. By choosing a BI cutoff score of ≥90 (sensitivity 90.7%; specificity 88.1%), the false favorable outcome rate could be reduced to 9.3%, whereas the false unfavorable outcome rate would increase to 11.9%. Minimizing false favorable and false unfavorable outcome rates could decrease unnecessary heterogeneity of outcome in acute stroke trials.

Compared with our results, Celani et al found that BI >90 (sensitivity 98%; specificity 97%) was a pivotal score for which patients did not require help from another person for everyday activities.\textsuperscript{14} Kay et al concluded that BI ≤80 (sensitivity 94%; specificity 80%) was the optimal cutoff score for self-reported dependency.\textsuperscript{12} These cutoff scores differed with those of our study when dependency is considered to be mRS ≥2, for which the optimal BI score was <90. These differences may be explained by the subjectivity of self-reported dependency, which will be influenced by personal circumstances such as socioeconomic status and psychological factors.

The mRS cutoff scores were used as a reference to distinguish favorable from unfavorable outcome. Although this is actually not a “gold standard” for dichotomizing outcome, we think that the mRS is suitable for this purpose. First, the mRS is a clinically relevant scale, with 6 different easily understandable and well-defined grades. Second, the BI is highly correlated with the mRS;\textsuperscript{19} therefore, we can compare BI cutoff scores with the mRS. Third, the mRS measures global disability, whereas the BI scores only ADL.

A point of criticism is that there is only a 5-point difference between the optimal BI cutoff scores in mRS 1 and mRS 2. These BI scores are near the maximum score of the BI. This can be explained by the frequently observed ceiling effects of the BI.\textsuperscript{5,6,20} Weimar et al concluded that because of the ceiling effect, the mRS is preferable to the BI for defining outcome.\textsuperscript{5} Kwon et al showed that there was no significant difference in BI scores between mRS 0, mRS 1, and mRS 2 because of ceiling effects of the BI.\textsuperscript{19}

If the intention of a therapeutic intervention is to obtain excellent recovery after stroke, which could be defined as mRS ≤1, the corresponding BI cutoff score was ≥95, according to our results. There is consensus that mRS ≤2 reflects independence and mRS >2 implicates dependence.\textsuperscript{21} Our study showed that a BI score ≥90 is the optimal cutoff score in relation to mRS ≤2. In severe strokes, one could decide to choose mRS ≤3 and BI ≥75 as cut-off scores for favorable outcome. An example of stroke severity–related outcome has been mentioned by Adams et al.\textsuperscript{22} They used the mRS as primary end point, where mRS cutoff scores 0, ≤1, or ≥2 reflected favorable outcome, depending on the baseline National Institutes of Health Stroke Scale (NIHSS) score.

In conclusion, we determined the optimal corresponding BI and mRS cutoff scores: BI 95 for mRS 1, BI 90 for mRS 2, and BI 75 for mRS 3. We recommend the use of this/these BI
cutoff score(s) with the corresponding mRS score(s) for future acute stroke trials in which BI and mRS scores dichotomize favorable and unfavorable outcome.

Acknowledgments
This study was supported by a grant from the Catharina Heerdnt Foundation.

References
Optimizing Cutoff Scores for the Barthel Index and the Modified Rankin Scale for Defining Outcome in Acute Stroke Trials
Maarten Uyttenboogaart, Roy E. Stewart, Patrick C.A.J. Vroomen, Jacques De Keyser and Gert-Jan Luijckx

Stroke. 2005;36:1984-1987; originally published online August 4, 2005;
doi: 10.1161/01.STR.0000177872.87960.61
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2005 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/36/9/1984

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/