Quantifying the Value of Stroke Disability Outcomes
WHO Global Burden of Disease Project Disability Weights for Each Level of the Modified Rankin Scale

Keun-Sik Hong, MD, PhD; Jeffrey L. Saver, MD

Background and Purpose—The modified Rankin Scale (mRS) categorizes poststroke disability among 7 broad, ordinal grades, but the interval distances between these levels are spaced along the disability spectrum have not been previously investigated.

Methods—We used the person trade-off procedure developed by the World Health Organization Global Burden of Disease Project (WHO-GBDP) to generate disability weights (DWs) ranging from 0 (normal) to 1 (dead) for each of 7 mRS grades. The ratings of an international, 9-member panel of stroke experts were combined by a modified Delphi process.

Results—DWs (95% CI) were 0 for mRS 0, 0.046 (0.004 to 0.088) for mRS 1, 0.212 (0.175 to 0.250) for mRS 2, 0.331 (0.292 to 0.371) for mRS 3, 0.652 (0.625 to 0.678) for mRS 4, 0.944 (0.873 to 1.015) for mRS 5, and 1.0 for mRS 6. DWs of adjacent mRS levels were significantly different (P<0.001 for all). Coefficients of variation showed a high degree of consensus for DWs among panel members. DWs placed each of the 5 intermediate mRS states in different disability class levels of the WHO-GBDP anchor conditions and identified natural clusters to use when reducing the mRS to fewer categories.

Conclusions—Formal DW assignment confirms that the mRS is an ordered but unequally spaced scale. The availability of DWs for each mRS level now permits direct comparison of each poststroke outcome state with the outcomes of hundreds of other diseases in the WHO-GBDP and the expression of stroke burden in different populations by using the uniform metric of disability-adjusted life-years lost. (Stroke. 2009;40:3828-3833.)

Key Words: disability weight • person trade-off • modified Rankin Scale • stroke outcome

Residual disability among stroke survivors spans a wide functional range. A variety of ordinal scales have been developed to classify patients along this spectrum, arraying individuals among a few broad, ordered levels. The most comprehensive and widely used ordinal poststroke global disability scale is the modified Rankin Scale (mRS), which assigns stroke survivors among 7 levels, including the extremes of normal and dead and 5 intermediate degrees of disability.

However, although ordinal scales are substantially more informative than binary-outcome measures, they are not as sensitive to change as are continuous outcome measures. Moreover, unlike interval scales, in which the distance between each scale step is uniform, ordinal scales array patients among unevenly spaced steps along the outcome range. Without knowing how far apart each step of an ordinal scale is along the disability spectrum, it is difficult to assess and compare therapies that exert differential beneficial and/or harmful effects at various scale state transitions.

Several methods have been developed to quantify how far apart on a health outcome dimension are the ranked levels of an ordinal scale. Most studies have used patient-preference techniques, including standard gamble, time trade-off, and visual analog scale methods. These techniques elicit utilities, how much value patients or healthy individuals place on specified outcomes,1-3 and permit analysis of outcomes in terms of quality-adjusted life-years (QALYs) gained or lost. However, when ratings are required for multiple rather than a single outcome state, these methods have the substantial weakness that patients and healthy individuals have limited, direct experience of different diseases and disease severities. Patients experience disease in depth, but not in breadth.

Using as raters physicians and allied health professionals, who have direct encounters with multiple diseases and disease severities, overcomes the unfamiliarity problem of patient-based quality-weighting techniques. Healthcare professionals observe disease both in breadth and in depth, albeit externally.4,5 Accordingly, tasked with deriving health outcome state ratings for all human diseases, the World Health Organization Global Burden of Disease Project (WHO-GBDP) developed the person trade-off (PTO) method.6 In the PTO framework, health professionals judge health states from...
a broad public health viewpoint that ensures equity across health states. This technique yields disability weights (DWs) for each disease, permitting analysis of treatment and policy outcomes in terms of disability-adjusted life-years (DALYs) gained or lost. The WHO’s PTO approach requires raters to undertake interpersonal comparisons of utility for individuals with different disease states, thus providing greater content validity to the resulting relative weights. In addition, the approach minimizes the confounders arising from adjustment to disease bias and lack of knowledge of a health state. The WHO’s DALY analyses have clarified understanding of the relative contribution of individual diseases to the total burden of disease in developed and underdeveloped regions and worldwide.

The PTO method can facilitate comparison of health outcomes not only across different diseases but also across varying severities of 1 disease. Consequently, the PTO method is well suited to assign quantified DWs for each of the ordinal levels of poststroke outcome scales.

Methods
We convened an international expert panel of stroke neurologists (7), a rehabilitation neurologist (1), and a double-board–certified emergency physician and neurologist (1) from the United States (4), South Korea (3), mainland China (1), and Taiwan (1).

DWs range from 0 (normal health without disability) to 1.0 (dead or as bad as being dead). DWs for mRS categories 0 and 6 were assigned 0 and 1.0, respectively. To determine the DWs for the 5 intermediate mRS categories, the panel members each individually performed the WHO-GBDP PTO1 exercise, in which they were asked to assume they were allocating health system resources and had to decide what number would make them indifferent between the choices of extending the lives of 100 healthy persons for 1 year versus extending the lives of N persons with the mRS health state for 1 year. As a reference framework during this task, the panel members were provided with the WHO-GBDP framework table, which displays 7 disability classes (DCs) and 22 anchoring example conditions (Table 1).

In the first round of the modified Delphi process, each expert individually provided ratings for the 5 intermediate mRS levels. The panel then met as a group and discussed each mRS level in turn. At the start of each level discussion, each panelist received a report showing (1) the group’s first-round mean and median PTO and disability class values for that mRS level, (2) the PTO and disability class values given by all respondents (anonymized) in the first round, and (3) the panelist’s own individual PTO rating in the first round. The group discussed the most appropriate PTO value for each of the 5 mRS states, with the moderator emphasizing full participation of all panel members, respect for all opinions, and continuation of discussion until distinctive thematic considerations were no longer elicited. Each panel member then again independently and confidentially assigned a PTO value for the discussed mRS state. The coefficient of variation (CV) was calculated to determine the need for additional Delphi rounds, with CV ≤0.5 indicating achievement of substantial consensus.

PTO values were converted to DWs by the formula $DW = \frac{1}{PTO}$. Because PTO values are exponentially distributed, their central tendency was assessed by calculating the geometric mean and 95% CI. A few experts assigned the mRS 5 state a PTO value of infinity (equal to death). In these instances, the extreme PTO value of $1\times10^6$ was substituted to permit calculation of geometric means. DWs are linearly distributed and were described by linear means and 95% CIs.

Normalization to Existing WHO-GBDP Rating for Poststroke Disability
The WHO-GBDP has generated DWs for acute stroke (0.920) and chronic poststroke disability (0.266). Although these ratings appropriately distinguish between acute and chronic poststroke health states, they fail to distinguish among chronic poststroke states of varying severity. The multispecialty WHO panel was well suited to compare disability across diverse organ-specific conditions but not to make fine distinctions among different stroke outcomes. In contrast, the PTO ratings of our stroke expert panel may be expected to be more reliable in delineating the relative value of different stroke outcomes compared with each other and less precise in determining the relative position of the disease category of stroke compared with hundreds of other, especially nonneurologic, diseases. To compare the DW ratings of the 5 intermediate mRS ranks generated by our stroke expert panel with the unitary chronic poststroke DW rating of the WHO-GBDP, we generated a composite DW for all stroke survivors from the DWs of individual mRS levels by using frequency-weighted averaging. For each mRS level, the panel-assigned DW was multiplied by the frequency of occurrence of that level among stroke survivors. To make the panel’s DWs for individual mRS levels fully consistent with the WHO-GBDP framework, an adjustment formula was derived to normalize the DWs for each mRS level to the WHO-GBDP unitary DW for all stroke survivors (supplemental mathematical Appendix available online at http://stroke.ahajournals.org).

Statistical Analyses
The Mann-Whitney test was used to compare the DWs of adjacent mRS states and DWs between the US and Asian panel members. The Wilcoxon signed-rank test was used to compare first- and second-round DW CVs.

Results
The mean DWs and geometric mean PTO values yielded by the panel members’ PTO ratings are shown in Table 2. All 5 intermediate disability states were assigned DW values consonant with their rank order on the mRS. DWs of adjacent mRS states were significantly different for all levels ($P<0.0001$). PTO values placed each of the 5 intermediate mRS states in different disability class levels of the WHO-GBDP anchor conditions, with mRS 1 assigned to WHO disability class (WHO-DC) 1, mRS 2 to WHO-DC 3, mRS 3

### Table 1. WHO-GBDP: DCs and PTO Numbers of Disabled Individuals of 22 Indicator Conditions

<table>
<thead>
<tr>
<th>DC</th>
<th>PTO No. of Disabled Individuals</th>
<th>Indicator Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000–1020</td>
<td>Vitiligo on face; weight for height &lt;2 SDs</td>
</tr>
<tr>
<td>2</td>
<td>1021–1136</td>
<td>Watery diarrhea; severe sore throat; severe anemia</td>
</tr>
<tr>
<td>3</td>
<td>1137–1316</td>
<td>Radius fracture in a stiff cast; infertility; erectile dysfunction; rheumatoid arthritis; angina</td>
</tr>
<tr>
<td>4</td>
<td>1317–1562</td>
<td>Below-the-knee amputation; complete deafness</td>
</tr>
<tr>
<td>5</td>
<td>1563–2000</td>
<td>Rectovaginal fistula; mild mental retardation; Down syndrome</td>
</tr>
<tr>
<td>6</td>
<td>2001–3333</td>
<td>Unipolar major depression; complete blindness; paraplegia</td>
</tr>
<tr>
<td>7</td>
<td>3334–infinity</td>
<td>Continuous psychosis; dementia; quadriplegia</td>
</tr>
</tbody>
</table>
to WHO-DC 4, mRS 4 to WHO-DC 6, and mRS 5 to WHO-DC 7.

For all mRS states, 95% CIs of DWs were narrow, <0.15 (Table 2). In the first round, the CV values for the DWs were <0.5 for all mRS ranks except for mRS 1 (Table 3). Inspection of individual rater values (supplemental Table I, available online at http://stroke.ahajournals.org) shows actual good agreement for DWs for mRS 1. The extremely low value range of the DWs of mRS 1 magnifies the relative value of small, absolute differences. The CVs decreased in the second Delphi round compared with those in the first round (P < 0.043). Because the overall CVs showed a high degree of consensus, further rounds were not necessary. Panel members from the United States and Asia showed no difference in DW ratings (P > 0.2 for all DWs).

The distance between 1 step mRS increments on the DW scale varied, confirming that the mRS is not an interval scale with equally spaced steps, but rather a rank-order scale with unequally spaced steps (the Figure). The closest mRS levels were mRS 0 and mRS 1, separated by 0.046; furthest apart were mRS 3 and mRS 4, separated by 0.321, a 7-fold difference. Modified Rankin Scale steps clustered among 4 groups along the DW scale: no disability, with or without symptoms=mRS 0 to 1; mild to moderate disability=mRS 2 to 3; severe disability=mRS 4; and extreme disability or death=mRS 5 to 6.

A MEDLINE review identified 6 studies in broad populations measuring the frequency of each mRS level among stroke survivors, 2 studies each for ischemic stroke (IS), intracerebral hemorrhage (ICH), and subarachnoid hemorrhage (SAH) (supplemental References available online at http://stroke.ahajournals.org). Because each study was performed among not fully comparable race-ethnic groups, the mRS frequencies for each stroke subtype were derived by averaging at the study level, rather than by weighting for sample size. Similarly, MEDLINE review identified 2 population-based studies reporting the incidence of different stroke subtypes, 1 from the Cincinnati/Northern Kentucky region in the United States and 1 from the Changsha region in China.14,15 The frequency of ICH was substantially higher among the Asian than the North American population. Based on the global perspective of the WHO-GBDP, these frequencies were averaged at the study level. This yielded incidence frequencies of stroke subtypes of IS=64.3%, ICH=29.8%, and SAH=2.1%. The relative frequency of each stroke subtype among all stroke survivors was derived by correcting for differential stroke subtype 30-day case-fatality rates (10.2% for IS, 37.6% for ICH, and 31.3% for SAH) to yield frequencies of each stroke subtype among stroke survivors of 74.3% IS, 23.9% ICH, and 1.9% SAH. The frequency-weighted average DWs for survivors of stroke subtypes were 0.223 for IS, 0.329 for ICH, and 0.209 for SAH (Table 4).

The frequency-weighted average DW for all stroke survivors was 0.248, which is close to the chronic poststroke DW of the WHO-GBDP multispecialty panel (0.266).13 Adjusted individual mRS DWs that normalize the stroke expert panel-generated DW for mRS level X, 1.052 is the adjustment factor, and 0.005 is the y-axis intercept.

**Table 2. PTO and DW for Each mRS Status**

<table>
<thead>
<tr>
<th>mRS 1</th>
<th>mRS 2</th>
<th>mRS 3</th>
<th>mRS 4</th>
<th>mRS 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geometric mean</td>
<td>1050</td>
<td>1272</td>
<td>1500</td>
<td>2881</td>
</tr>
<tr>
<td>Geometric 95% CI</td>
<td>1003–1099</td>
<td>1211–1335</td>
<td>1415–1590</td>
<td>2676–3102</td>
</tr>
<tr>
<td>DW Mean</td>
<td>0.046</td>
<td>0.212</td>
<td>0.331</td>
<td>0.652</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.004–0.088</td>
<td>0.175–0.250</td>
<td>0.292–0.371</td>
<td>0.625–0.678</td>
</tr>
<tr>
<td>Adjusted DW</td>
<td>0.053</td>
<td>0.228</td>
<td>0.353</td>
<td>0.691</td>
</tr>
</tbody>
</table>

**Table 3. CVs of DWs for Each mRS Status in the First and Second Rounds**

<table>
<thead>
<tr>
<th>Round</th>
<th>mRS 1</th>
<th>mRS 2</th>
<th>mRS 3</th>
<th>mRS 4</th>
<th>mRS 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.40</td>
<td>0.41</td>
<td>0.25</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>2</td>
<td>1.19</td>
<td>0.23</td>
<td>0.16</td>
<td>0.05</td>
<td>0.10</td>
</tr>
</tbody>
</table>

P = 0.043 for comparison of the CVs of DWs between the first and second Delphi round (Wilcoxon signed-rank test).

**Discussion**

This study is the first to derive DWs for each level of the mRS of poststroke global disability. The internal validity of the study is supported by the low CVs among the raters and the derived DWs conforming to the original rank order of the mRS. The
external validity is supported by consistency with (1) 2 prior DW studies of broader poststroke categories, (2) a survey study of physician intuitions regarding differences in mRS levels, (3) patient-derived quality-of-life (QOL) utility ratings of poststroke states, and (4) concurrent validity studies of complementary functional measures associated with different mRS levels.

Our results concur with and extend 2 prior studies of DWs for poststroke health states.10,13 The composite DW for all stroke survivors yielded by frequency-weighting individual DWs for each mRS level in the current study (0.248) is quite close to the unitary DW for the chronic poststroke state derived by the WHO-GBDP multispecialty panel (0.266).13 As a result, calibrating the fine-grained state distinctions made by stroke experts to the broad global state assessment made by a multispecialty panel required minimal adjustment (Table 2).

A Dutch study used the PTO to derive DWs for 3 categories of poststroke impairment: severe, moderate, and mild.10 The DW of 0.94 for mRS level 5 in this study is consistent with the DW of 0.8 to 1.0 for severe poststroke impairment in the Dutch study, and the DW of 0.65 for mRS level 4 in the present study comports with the Dutch DW of 0.5 to 0.65 for moderate poststroke impairment. In contrast, the DWs of 0.05, 0.21, and 0.33 for mRS levels 1, 2, and 3 in the current study are generally lower than the Dutch DW for mild poststroke impairment of 0.3 to 0.4. This may reflect differences in the specialty composition of the expert panels, as the general practitioners in the Dutch study may have less familiarity and experience with mild poststroke impairments than the stroke physicians in our study. Cultural difference regarding equity, an important consideration in determining PTO values, might also account for the differences in the DW ratings. However, because our panelists were composed of multinational experts (although none of European countries), the cultural difference might not be an important determinant factor. The current study goes beyond prior DW investigations to provide specific DWs for a more granular series of poststroke states, yielding valuations with greater sensitivity to change over the full stroke outcome range.

The findings of the current study largely conform to a prior, informal questionnaire study delineating stroke physician valuation of individual mRS levels.16 The survey found that stroke physicians value health state transitions across all mRS levels, with the frequent exception of transitions from extreme disability (mRS 5) to death (mRS 6), which are often collapsed into a single worst-outcome category.17 With the use of PTO methodology, the current study formally demonstrated a monotonic, ordered separation of DWs for all mRS levels but with the smallest adjusted DW separation between mRS 5 and 6.

No prior study has derived patient-preference QOL utility ratings for each mRS level. However, many studies of broader poststroke categories have been undertaken, typically delineating separate utilities for 2 (minor and major stroke) or 3 (minor, moderate, and major stroke) poststroke states. To compare these utility values with DWs, it is important to recognize that QOL utility analyses assign a score of 0 for being as bad as dead and 1 for being as good as normal, and thus yield inverse values compared with DWs. To permit direct comparison, QOL utility values may be converted to quality weights (QWs) by the formula $QW = 1 - utility$. A systematic review of all studies deriving QOL utility ratings by patient time trade-off preferences for poststroke states found that, to minor stroke, stroke survivors assigned a QW of 0.28 and nonstroke patients with vascular risk factors a QW of 0.45, whereas to major stroke, stroke survivors assigned a QW of 0.59 and nonstroke patients with vascular risk factors a QW of 0.74.18 The DWs in this study for the mRS levels comprising minor stroke (mRS 1, 2, and possibly 3) and major stroke (4, 5, and possibly 3) are roughly consonant with the QWs previously derived from patient time trade-off techniques.

It is noteworthy that the physician-derived DWs correspond closely with the stroke patient–derived QWs, in contrast with the QWs derived from nonstroke patients, which are substantially more severe. It is a common finding that patients rate their disability less than do nonpatients.18 This pattern has been attributed to several factors, including adaptation to disease by affected patients (“adjustment to disease bias”), and nonpatients’ fear of unaccustomed conditions (“unfamiliarity with disease bias”). We hypothesize that physician-derived ratings resemble affected patient–derived ratings because physicians are knowledgeable regarding the disease state and have observed the substantial capacities of their patients to adjust to physical challenge.

### Table 4. Frequencies of mRS Outcomes Among Stroke Survivors and Resulting Frequency-Weighted DWs for Stroke

<table>
<thead>
<tr>
<th></th>
<th>IS</th>
<th>ICH</th>
<th>SAH</th>
<th>All Stroke*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IS</td>
<td>ICH</td>
<td>SAH</td>
<td>All Stroke*</td>
</tr>
<tr>
<td>mRS 0</td>
<td>23.3</td>
<td>30.2</td>
<td>26.7</td>
<td>0.228</td>
</tr>
<tr>
<td>mRS 1</td>
<td>30.2</td>
<td>24.2</td>
<td>27.2</td>
<td>0.218</td>
</tr>
<tr>
<td>mRS 2</td>
<td>17.4</td>
<td>16.0</td>
<td>16.7</td>
<td>0.223</td>
</tr>
<tr>
<td>mRS 3</td>
<td>9.3</td>
<td>13.2</td>
<td>11.2</td>
<td>0.354</td>
</tr>
<tr>
<td>mRS 4</td>
<td>14.0</td>
<td>9.3</td>
<td>11.6</td>
<td>0.231</td>
</tr>
<tr>
<td>mRS 5</td>
<td>5.8</td>
<td>7.6</td>
<td>6.6</td>
<td>0.218</td>
</tr>
<tr>
<td>Composite DW</td>
<td>0.228</td>
<td>0.218</td>
<td>0.223</td>
<td>0.209</td>
</tr>
</tbody>
</table>

References of the individual studies are listed in the supplemental references, available online at http://stroke.ahajournals.org.
This study’s findings of distinctive, separate DWs for each mRS level are also correlated with prior studies of the clinical and economic course of patients at different mRS grades.\textsuperscript{19,20} The proportion of the first 3 poststroke months spent at home decreases as mRS level increases, long-term life expectancy decreases as mRS level increases, and cost of care increases as mRS level increases from 0 to 5.

That adjacent mRS levels are unequally spaced along the disability spectrum has important implications for interpreting the results of stroke treatment in clinical trials and in health policy models. In the past, when evaluating socioeconomic consequences of trial results and policy changes, an equal-interval distance between mRS levels had to be assumed.\textsuperscript{21} Now, actual DW value distances can be used, enabling more accurate assessment of the human impact of stroke treatments and public health policies.

Sometimes it is desirable to reduce the full 7-level mRS to fewer categories to simplify analysis. Hitherto, the assortment of the 7 mRS levels into fewer categories was performed in an intuitive, somewhat arbitrary, manner. Supraordinate category formation may now be guided by the natural clustering of the formally derived, quantitative DWs observed in this study. For example, thrombolysis trialists have reduced the mRS to 4 categories: mRS 0 to 1, mRS 2 to 3, mRS 4 to 5, and mRS 6.\textsuperscript{22} However, the clustering of DWs derived in this study indicates that a more natural tetrachotomization of the Rankin Scale, combining mRS levels with similar DWs, would be mRS 0 to 1, mRS 2 to 3, mRS 4, and mRS 5 to 6. Another approach to tetrachotomization was taken by the CLOTBUST trialists, who reduced the mRS to 0 to 1, 2, 3 to 5, and 6.\textsuperscript{23} The DWs here derived suggest that the CLOTBUST approach yields a distorted tetrachotomy, with the single mRS 3 to 5 category spanning 70% of the disability spectrum and the remaining 3 categories subdividing only the remaining 30%. This might be useful for a clinical trial aimed at detecting only transitions to, from, and among a range of good outcomes but will miss important changes in outcome in the moderate to severe range. An additional approach in stroke trial analysis is to reduce the mRS to 6 levels, rather than 7, by collapsing mRS 5 and mRS 6 into a single worst-outcome category.\textsuperscript{17} This study supports this hexachotomization approach, indicating little difference along the disability dimension among mRS 5 and 6 outcomes.

This study has limitations. The expert panel included stroke physicians from North America and Asia but not other regions of the world. The panel included representatives of 3 specialties/subspecialties (stroke neurology, rehabilitation neurology, emergency medicine) but not additional specialties involved in stroke care, nor nurses or allied health personnel. Some panelists who rated mRS 5 equivalent to death may have actually considered extreme poststroke disability to be worse than death, but the PTO task does not permit scoring an outcome state as worse than death.

Stroke is the only disease that is both a leading cause of death and a leading cause of disability worldwide. Although stroke death rates are inherently easy to enumerate, the total burden of poststroke disability has been challenging to quantify. The availability of DWs for each mRS level now permits direct comparison of each poststroke outcome state with the outcomes of hundreds of other diseases in the WHO-GBDP and the expression of stroke burden in different populations by using the uniform metric of DALYs lost. Because DALYs are a continuous outcome measure, stroke treatments and health policy effects can now be examined not only with dichotomized and ordinal statistical methods but also with the often more powerful methods appropriate to continuous data.

**Acknowledgments**

The authors are grateful to the members of the expert panel: David Alexander, MD; Si-Ryung Han, MD; Keun-Sik Hong, MD; Meng Lee, MD; David S. Liebeskind, MD; Jaseong Koo, MD; Jeffrey L. Saver, MD; Sidney Starkman, MD; and Liangfu Zhu, MD.

**Sources of Funding**

This work was supported by a grant (A060171) of the Korea Health 21 R&D project, Ministry of Health, Welfare and Family Affairs, Republic of Korea (K.-S.H.), and NIH-NINDS Award P50 NS044378 (J.L.S.).

**Disclosures**

None.

**References**


3832 *Stroke* December 2009