



University  
of Glasgow

Quinn, T.J., Lees, K.R. , Hardemark, H.G., Dawson, J. , and Walters, M.R. (2007) *Initial experience of a digital training resource for modified Rankin scale assessment in clinical trials.* Stroke, 38 (8). pp. 2257-2261.  
ISSN 0039-2499

<http://eprints.gla.ac.uk/21004/>

Deposited on: 19 January 2012

# Stroke

American Stroke  
Association<sup>SM</sup>

JOURNAL OF THE AMERICAN HEART ASSOCIATION

A Division of American  
Heart Association



## **Initial Experience of a Digital Training Resource for Modified Rankin Scale Assessment in Clinical Trials**

Terence J. Quinn, Kennedy R. Lees, Hans-Goran Hardemark, Jesse Dawson and Matthew R. Walters

*Stroke* 2007, 38:2257-2261: originally published online June 28, 2007

doi: 10.1161/STROKEAHA.106.480723

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75214  
Copyright © 2007 American Heart Association. 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/38/8/2257>

Subscriptions: Information about subscribing to *Stroke* is online at  
<http://stroke.ahajournals.org/subscriptions/>

Permissions: Permissions & Rights Desk, Lippincott Williams & Wilkins, a division of Wolters Kluwer Health, 351 West Camden Street, Baltimore, MD 21202-2436. Phone: 410-528-4050. Fax: 410-528-8550. E-mail:  
[journalpermissions@lww.com](mailto:journalpermissions@lww.com)

Reprints: Information about reprints can be found online at  
<http://www.lww.com/reprints>

# Initial Experience of a Digital Training Resource for Modified Rankin Scale Assessment in Clinical Trials

Terence J. Quinn, MRCP; Kennedy R. Lees, MD; Hans-Goran Hardemark, PhD;  
Jesse Dawson, MRCP; Matthew R. Walters, MD

**Background and Purpose**—The modified Rankin Scale (mRS) is the preferred measure of disability in cerebrovascular clinical trials, but its value is restricted by interobserver variability. Poor reliability reduces the statistical power of clinical trials and leads to underestimation of effect size. Strategies to improve mRS grading are required. Video training has previously improved application of the National Institutes of Health Stroke Scale in clinical research. We developed an mRS training resource in an attempt to minimize interobserver variability.

**Methods**—We produced a complete training resource comprising an instructional DVD with accompanying written materials and assessment recordings of patient interviews. Formal assessment of training involved grading of real-life cases. Results of initial training and recertification were collected centrally and scored.

**Results**—Data from 1564 assessments are presented. The majority of assessors were participating in 2 large prospective clinical stroke trials. Assessors represented a mixed group of disciplines and nationalities. After training, most trainees (90%) achieved certification in mRS assessment. The majority (85%) of investigators who did not reach an acceptable score on initial testing achieved certification after further exposure to the package.

**Conclusions**—Mass training in mRS assessment for clinical trials is possible. We outline the development of a video-based training package, including technical issues, patient selection procedures, and methods of scoring and assessment. Certification results suggest that use of the resource can improve mRS grading. Acceptability of the training has been demonstrated by its successful use in 2 international acute stroke trials, SAINT 1 and CHANT. (*Stroke*. 2007;38:2257-2261.)

**Key Words:** cerebrovascular accident ■ medical education ■ modified Rankin Scale  
■ outcome assessment ■ stroke treatment ■ video recording

Assessment of recovery from stroke requires a valid and reliable outcome measure. The modified Rankin Scale (mRS) is now the preferred disability end point for clinical trials<sup>1</sup> and has been used in a number of interventional studies.<sup>2,3</sup> It is an ordinal hierarchical scale incorporating 6 categories that describe the range of disability encountered after stroke.<sup>4</sup> Although widely used, the mRS has been criticized for its relative lack of structure<sup>5</sup>: Rankin grades encompass a broad range of potential outcomes, and boundaries between grades are poorly defined relative to other outcome assessment instruments. This lack of structure is reflected by the high interobserver variability seen when the mRS is applied. Single-center studies typically give weighted  $\kappa$  statistics between 0.7 and 0.8,<sup>6</sup> whereas a 3-site comparison designed to simulate a multicenter trial reported an unweighted  $\kappa$  statistic of only 0.25 with the standard mRS assessment.<sup>7</sup>

Various strategies have been proposed to improve reliability, including video recording<sup>8</sup> and use of a structured interview.<sup>7,9</sup> Improving reliability of the mRS is of more than

clinimetric interest; interobserver variability will increase the risk of end-point misclassification, which can introduce bias and affect type II error rates.<sup>10</sup> Estimated effect size will be reduced. It has been argued that the statistical underpowering has contributed to the lack of significant treatment effects seen in recent acute stroke trials.<sup>11</sup>

Use of a training resource to improve consistency in the application of the mRS makes intuitive sense.<sup>12</sup> In the original work describing interobserver variation in mRS scoring, the authors commented that improvements could be achieved if observers were afforded the chance to practice use of the scale, but that "... such training is hardly realistic in the context of a multicenter trial."<sup>4</sup> Recent improvements in audiovisual technology mean mass training across a number of centers is now feasible.

A variety of potential formats are available for training. A stand-alone package would allow better standardization and dissemination than a lecture-based series, whereas a program incorporating "live" assessment of real patients is preferable to purely text-based instruction. The former approach is not

Received December 20, 2006; final revision received January 31, 2007; accepted February 20, 2007.

From the Division of Cardiovascular and Medical Sciences (T.J.Q., K.R.L., J.D., M.R.W.), University of Glasgow, Glasgow, Scotland, and AstraZeneca R&D Södertälje (H.G.H.), Medical Neuroscience, Södertälje, Sweden.

Correspondence to Dr Terence J. Quinn, Gardiner Institute of Cardiovascular and Medical Sciences, Western Infirmary, Glasgow G116NT, UK. E-mail tjq1t@clinmed.gla.ac.uk

© 2007 American Heart Association, Inc.

Stroke is available at <http://www.strokeaha.org>

DOI: 10.1161/STROKEAHA.106.480723

without precedent. A well-validated video-based training program and certification procedure exist for the National Institutes of Health Stroke Scale (NIHSS), and it is now a requirement that clinical trial investigators complete this training and undergo certification.<sup>13</sup> The strengths of the NIHSS program, in particular its mix of didactic teaching, video explanation, and assessment procedure, could be easily applied to the mRS.

We developed a training DVD and accompanying explanatory booklet, which include recordings of real Rankin assessments, certification cases, and further recertification cases. This has been used successfully in 2 large-scale clinical stroke trials.<sup>14,15</sup> Although brief reference to certification scoring has been made in a previous review,<sup>16</sup> to date there has been no detailed description of the training package, its development, and the initial experience of its use. We present this here along with the results of the certification program.

## Materials and Methods

### Development of the Training Package: Audiovisual Issues

A criticism of the early NIHSS video training was poor image quality, an important consideration, as subtle clinical signs may have a substantial effect on final scoring.<sup>16</sup> In an interview-based assessment such as the mRS, high-quality sound recording is of equal importance. We enlisted expert technical help (Media Services Department, University of Glasgow). A recording studio was used and, where patient disability made travel to the studio inappropriate, the Media Services mobile crew filmed in the hospital.

We chose a DVD-based format. This reflects the gradual replacement of conventional VHS recordings with DVD and the improved clarity of sound and vision afforded by it. Digital recording ensures optimal quality, even after mass reproduction. Finally, this allows for ease of transfer to a Web-based server. Internet-based NIHSS training has facilitated cost-effective global dissemination of the training package.

### Patient Selection

The stroke liaison team based in the Western Infirmary, Glasgow, selected suitable patients from recent admissions to the acute stroke unit or local inpatient rehabilitation facility. The stroke unit accepts all patients presenting within 72 hours of onset of suspected stroke, irrespective of age or severity of the neurologic deficit; thus, a cohort of patients with varying degrees of disability and background comorbidity was available. Our intention was to include at least 1 patient from each potential mRS category. Thirteen patients were selected, of whom 4 were designated as training cases, 5 were used for initial assessment, and 4 were used for recertification. To reflect clinical practice, in 1 case disability was sufficiently severe that answers were provided by a caregiver. Although the final selection of patients was a sample of convenience, they were thought to be representative of a "real-life" cross section of poststroke outcomes and to be suitable for inclusion in the training package. We obtained a consent for videotaping and use for training and research purposes from all patients, in line with national and local protocols.

For the training component of the package, 2 patients with easily categorized disability were chosen as initial "introductory" cases. The remaining 2 training cases were chosen to highlight perceived problem areas in the application of the scale. Each of the training cases was followed by an explanatory discussion of the correct mRS score and the rationale for this grading. An accompanying booklet gave background information on the general principles of mRS scoring, including detailed definitions of the categories and discussion of what is considered to be best practice in the application of the scale. In formulating this advice, we made reference to the original description of the mRS and to recent work by means of a structured

**TABLE 1. Scoring System for Certification Process**

Grading	Certification
All 5 cases correct	Pass (qualified 5)
4 correct cases and 1 acceptable	Pass (qualified 4)
3 correct cases and 2 acceptable	Pass (qualified 3)
Any other combination	Not qualified

interview.<sup>4</sup> To minimize potential language problems, we provided a transcript of the text with translation into local language. Fully translated training packages, with native speakers overdubbing the interview, have been made available for Spanish, Portuguese, and Italian researchers; a subtitled Chinese version has also been produced. Assessment for certification was performed in a variety of settings, including individual viewing of the cases, group viewing within a center, and supervised group viewing sessions at formal training meetings.

### Recording and Scoring of the Assessments

Recordings of the interviews were analyzed and scored independently by 2 observers who were both experienced in mRS grading (K.R.L. and H.G.H.). No attempt was made to "script" the mRS assessments, and there was little postinterview editing. During a pilot run of the study, we immediately identified that some of the answers given by patients were ambiguous, and in at least 2 of the cases, debate arose as to the most appropriate category. However, a decision was made to keep the complete interviews as recorded; it was thought that scenarios artificially scripted to fit an mRS grade neatly would not have adequately prepared assessors for the difficulties inherent in grading real patients.

Scoring for the assessment component of the package took account of those patients who did not unequivocally fit a single grade. A final decision on correct grading was made by K.R.L. and H.G.H., supplemented by analysis of the results from an international pilot study involving 100 participants. We defined a correct grade as one assigned by both trainers and by >50% of trainees. A grade of an "acceptable" response was defined arbitrarily as one that was deemed by the trainers to be incorrect but that followed the basic scoring guidelines and had been assigned by a substantial minority (10% to 49%) of assessors in the pilot. Any grade offered by <10% of assessors or that clearly did not follow the accepted scoring was defined as "unacceptable." According to these scoring rules, any candidate who graded all cases correctly, including "acceptable" answers, was awarded certification (Table 1). This scoring system was developed so that certification was awarded only to those assessors who demonstrated a good knowledge of the mRS application but recognized that some merit should be given for "acceptable" but incorrect answers. To minimize variability, we instructed assessors to choose the more severe mRS grade when hesitating between 2 scores.

We awarded certificates of completion, along with separate confidential feedback on the actual score achieved, to all who achieved the target. Scores for individual patients were not released. We encouraged assessors who did not achieve certification on their first attempt to review the training material and resubmit an amended set of grades for the full set of certification cases. We gave no specific feedback on their errors.

## Results

Between March 2003 and May 2006, investigators from >25 countries submitted the end-of-training assessment. Data are available on 1800 assessments. The assessors were a mixed group of principal investigators and coinvestigators, study nurses, and research assistants (Table 2). The majority of respondents were part of the investigating team from countries involved in the Stroke-Acute Ischemic-NXY-059 Treatment (SAINT-I)<sup>14</sup> trial only or in both

**TABLE 2. mRS Gradings Submitted for Certification and Recertification, %\***

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
mRS 0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.1
mRS 1	0.0	0.9	26.5	2.7	0.0	0.0	0.0	0.0	2.2
mRS 2	0.7	55	72.6	39.2	0.0	1.5	2.3	0.1	40.4
mRS 3	6.3	44.1	0.7	58.1	0.0	3.7	90.4	1.7	56.2
mRS 4	92.1	0.0	0.0	0.0	14	93.8	7.3	97.4	1.0
mRS 5	0.9	0.0	0.0	0.0	86	0.9	0.0	0.6	0.0
Correct grade	4	2	2	3	5	4	3	4	3

\*To allow for future use of the assessment exercise, cases are not given in the order they appear on the DVD.

the SAINT-II and the Cerebral Hemorrhage and NXY-059 Treatment (CHANT) study.<sup>15</sup>

The correct mRS scores for the certification cases, along with the proportion of scores assigned by observers, are shown in Table 3. To allow continued use of the training resource, the “correct” scores for assessment and recertification were not made public and are shown in a sequence different from the cases on the DVD.

There was a spread of opinion on all of the cases, with submitted answers spanning 3 to 5 mRS grades for each answer. For 3 of the cases, the majority of respondents opted for the correct grading; in 2 of the cases, opinion was split, with a substantial proportion of assessors (39.2% and 40.4%) choosing a lower mRS grade than the correct answer. We accounted for this variation in opinion in the final scoring, and the lower grade was defined as an “acceptable” answer. Twenty-three assessors gave 2 scores, despite explicit instructions to choose the best single score. If the scores given were the correct and an acceptable score, then a “pass” was allowed; otherwise, the grading was considered invalid.

Percentages of respondents achieving acceptable scores for the certification assessment are given in Table 2. The majority of assessors (1464, 81.3%), achieved a “pass” on the certification exercise. However, only 38% of these individuals graded all of the 5 cases correctly. The remainder of the group comprised those assessors who wrongly assessed 1 or both of the previously described equivocal cases but whose assessment was still defined as acceptable. Of the 336 who did not achieve certification on their first attempt, 85% scored a “pass” on a second attempt.

Demographic data on assessors submitting for certification were collated. Results of the certification assessment are presented by training (Table 2) and by country of origin (Figure). We intended the recertification process to be undertaken 1 year after the initial training, as such full data on recertification are presently limited to 370 results. Of these,

only 6.5% assessors failed to achieve a satisfactory score (Table 3).

### Discussion

Consistency in grading of poststroke disability is crucial both in daily clinical work and in the context of a clinical trial. For trial purposes, consistency is more important than accuracy. The potential for significant variation in application of the mRS is now apparent.<sup>6</sup> We developed a digital training resource for mRS grading in an attempt to improve this situation. Our results show that mass training of observers in use of the mRS is achievable in the context of a clinical trial via the use of a novel DVD training package.

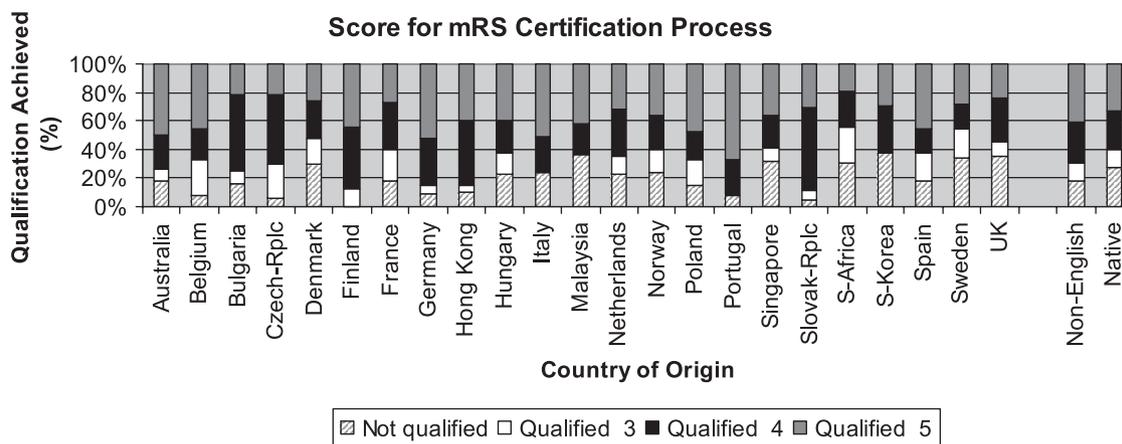
Several issues arose during development of the DVD that deserve comment. We used no specific criteria to select patients for the training or certification components of the package, although patients used for assessment were judged to be suitably taxing to allow a valid assessment of ability. A clustering of grades around the midrange was noted in the cases selected to be used for the certification process. It has been shown that clinicians are comfortable to assign grades at the extremes of the mRS, possibly because these grades are well-defined or because deviation can be in 1 direction only. Therefore, we decided to proceed to use this relatively biased sample.

Even with a training system, it may be impossible to completely remove interobserver variability for the “difficult” midrange grades. For assessment cases 3 and 9, respondents were almost evenly split between mRS grades 2 and 3. This is not a failure of the training, rather an example of the complexity of assigning fixed grades to real patients. We anticipate that those video cases that have divided opinion will be reviewed again by trainees, and discussion of these cases with other investigators will facilitate further improvement in mRS application. It is in the midrange of mRS grades that the reliability of the scale assumes the greatest impor-

**TABLE 3. Certification Scores by Background Training\***

Position	Not Qualified	Qualified 3	Qualified 4	Qualified 5	Total
Coinvestigator	125 (16.7%)	96 (12.9%)	224 (30.0%)	302 (40.4%)	747
Principal investigator	28 (17.6%)	18 (11.3%)	48 (30.2%)	65 (40.9%)	159
Study nurse	53 (25%)	29 (13.7%)	64 (30.2%)	66 (31.1%)	212
Other	130 (19.0%)	71 (10.4%)	228 (33.4%)	253 (37.2%)	682
Total	336	214	564	686	

\*For explanation of “Qualified 5, 4, 3,” see Table 1.



Certification scores by country. "Native" refers to native English-speaking countries; "non-English", to countries where English is not the first language. For explanation of "qualified 5, 4, 3," see Table 1.

tance; mRS-based outcomes are frequently dichotomized, with mRS  $\geq 4$  defining poor outcome (STICH trial)<sup>3</sup> or scores  $\leq 2$  defining good outcome (ECASS II).<sup>2</sup> Given the problems with midrange mRS grades highlighted by our data, increasing use of nondichotomized trial end points should be encouraged. To this end, the SAINT and CHANT trials examined changes in mRS across the continuum of possible grades.

It is difficult to adequately measure the "success" of a multidimensional intervention such as an educational resource. Analyzing only "pass" rates for the assessment is a relatively crude measure, as it is likely that even those viewers who failed the initial certification will have gained improved knowledge of the mRS scores and assessments. Despite this, the high rate of satisfactory scoring on the certification exercise is reassuring.

We gave assessors confidential feedback on their total score. This allowed all users to review the cases and perhaps to correct any grading errors. Data are not available for all of the second attempts at certification, which raises the possibility of sample bias, but the high "pass" rate seen provides further support for utility of the training package. Given that our training package is the only educational resource available for training, it is safe to assume that this improvement was achieved with no extra tuition other than repeated viewing of the package and knowledge of previous scoring.

Purists will argue that our data do not conclusively prove the benefit of the training package. The primary purpose of the package was to improve reliability of mRS grading in large clinical trials, so we did not design a "control" arm of assessors not exposed to training. Extrapolating evidence from the success of the NIHSS and other video certification schemes<sup>17</sup> suggests this was a logical and ethically sound decision. It is widely accepted that there is too little formal guidance on application of the mRS,<sup>18</sup> and few would argue against an attempt at formalizing its use. Anecdotally, feedback from participants has been uniformly positive, and there is little to suggest that exposure to the digital training worsens mRS grading or introduces systematic bias. Even if the training were to influence grading systematically, it could be argued that if all assessors were taught to grade in the same

fashion and in a manner that reflects the mRS categories, that this could only improve outcome assessments for trial purposes.

Pragmatic evidence of the utility of the training package comes from its application in the SAINT-I<sup>14</sup> and CHANT<sup>15</sup> studies. During conduct of these studies, >1500 investigators were trained. It cannot be proved that this improved the quality of end-point assessment, but it does demonstrate the feasibility of mass training. Given the inherent problems with use of the mRS, it is unlikely not to have helped. We believe that formal mRS training should be routine for all acute stroke trials.

Even with the improvements in mRS grading offered by use of the training package, there remains considerable scope to further reduce interobserver variability. The increased availability of affordable audiovisual equipment makes multicenter digital recording of mRS interviews feasible. Such an approach would allow further review of interviews or "off-line" assessment by other individuals or groups experienced in the use of the scale. We are currently piloting the use of such a scheme in our acute stroke unit.

It is recognized that some questions as to optimal delivery of the package remain unanswered, such as how best to address the issue of repeated failure to achieve certification and whether training should be performed alone or in a group setting with the opportunity to discuss the cases and content. Work is ongoing to answer these questions.<sup>17</sup> Already there is scope for further improvements; for instance, making the training available on the Internet could ease dissemination to the target audience.

We recognize that many stroke researchers will not have English as a first language. To facilitate improved mRS scoring internationally, foreign language versions of the training resource have been developed. It is interesting to note that the majority of non-English-speaking countries, including many that do not yet have a native language training package, actually achieved better scores on the certification exercise than did their UK counterparts.

Although accepted in the stroke literature from its inception, the mRS was not clinimetrically tested before its use.<sup>4</sup> The data being collated from the mRS certification process

provide a powerful tool for better definition of the properties of the scale. Analysis of interobserver and intraobserver variability of the scale with further subanalysis of individual components of the scale and relations to country of origin and level of training are ongoing.

We live in a digital age, and using the available technology to deliver educational resources makes scientific and economic sense. Strategies to improve the reliability of the mRS are needed, and it is likely that electronic dissemination of teaching material to participants in multicenter clinical trials will be widely used in the future. We have demonstrated that digital training in poststroke assessment is feasible and accepted by most potential assessors. Further work to quantify the potential impact of such training on the quality of future stroke trials is required.

### Acknowledgments

We are grateful to Colin Brierley, Nigel Hutchins, Barbara Farmer, and their team at University of Glasgow Media Services for assistance with video recordings. We would also like to thank Sarah Dorward, our stroke liaison sister, and all of the patients and staff who assisted in production of the training package. Finally, we acknowledge all of the SAINT-I and CHANT steering committee and investigators, Dr Algirdas Kakarieka, and the AstraZeneca monitors who helped in administering many of the assessments and all other researchers who have completed the certification exercise and/or provided comments on the training resource.

### Source of Funding

The development of this resource was partly supported by an educational grant from AstraZeneca; rights to use of the training material were retained by K.R.L. SAINT I and CHANT were sponsored by AstraZeneca, and NXY-059 is being developed by AstraZeneca under a license agreement with Renovis.

### Disclosures

K.R.L. was the international principal investigator for the GAIN International and SAINT I trials and chairs the steering committee for the CHANT and SAINT I and II trials. He has published data in support of the mRS as the optimal end point for acute stroke trials. He has received fees, expenses, and institutional grants relating to these and other trials from GlaxoSmithKline, AstraZeneca, and several other pharmaceutical companies that have been or are developing treatments for stroke. H.G.H. is an employee of AstraZeneca. K.R.L., M.W., J.D., and T.J.Q. have applied for academic grant support to continue work on developing stroke outcome assessments with the mRS.

### References

1. Roberts L, Counsell C. Assessment of clinical outcomes in acute stroke trials. *Stroke*. 1998;29:986–991.
2. Second European-Australasian Acute Stroke Study Investigators. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). *Lancet*. 1998;352:1245–1251.
3. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial. *Lancet*. 2005;365:387–397.
4. Van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJA, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988;19:604–607.
5. Wolfe CDA, Taube NA, Woodrow EJ, Burney PG. Assessment of scales of disability and handicap for stroke patients. *Stroke*. 1991;22:1242–1244.
6. Duncan PW, Jorgensen HS, Wade DT. Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice. *Stroke*. 2000;31:1429–1438.
7. Wilson JTL, Hareendran A, Hendry A, Potter J, Bone I, Muir KW. Reliability of the modified Rankin Scale across multiple raters: benefits of a structured interview. *Stroke*. 2005;36:777–781.
8. Shinohara Y, Minematsu K, Amano T, Ohashi Y. Modified Rankin Scale with expanded guidance scheme and interview questionnaire: interrater agreement and reproducibility of assessment. *Cerebrovasc Dis*. 2006;21:271–278.
9. Wilson JTL, Hareendran A, Grant M, Baird T, Schulz UGR, Muir KW, Bone I. Improving the assessment of outcomes in stroke: use of a structured interview to assign grades on the modified Rankin Scale. *Stroke*. 2002;33:2243–2246.
10. Choi SC, Clifton GL, Marmarou A. Misclassification and treatment effect on primary outcome measures in clinical trials of severe neurotrauma. *J Neurotrauma*. 2002;19:17–12.
11. Lees KR, Hankey GJ, Hacke W. Design of future acute-stroke treatment trials. *Lancet Neurol*. 2003;2:54–61.
12. Garraway WM, Akhtat AJ, Gore SM, Prescott RJ, Smith RG. Observer variation in the clinical assessment of stroke. *Age Ageing*. 1976;5:233–240.
13. Lyden P, Brott T, Tilley B, Welch KMA, Mascha EJ, Levine S, Haley EC, Grotta J, Marler J. NINDS TPA Stroke Study Group. Improved reliability of the NIH Stroke Scale using video training. *Stroke*. 1994;25:2220–2226.
14. Lees KR, Zivin JA, Ashwood T, Davalos A, Davis SM, Diener HC, Grotta J, Lyden P, Shuaib A, Hardemark HG, Wasiewski WW. Stroke-Acute Ischemic NXY Treatment (SAINT I) Trial Investigators. NXY-059 for acute ischemic stroke. *N Engl J Med*. 2006;354:588–600.
15. Lyden PD, Shuaib A, Lees KR, Davalos A, Davis SM, Diener H-C, Grotta JC, Ashwood TJ, Hardemark H-G, Svensson HH, Rodichok L, Wasiewski WW, Ahlberg G. Final results of CHANT: a study of the safety and tolerability of NXY-059 in intracerebral hemorrhage. *Stroke*. 2007;38:475.
16. Lees KR, Milia P. Halving effort in acute stroke trials. *Clin Exp Hypertens*. 2006;28:1–4.
17. Lyden P, Raman R, Liu L, Grotta J, Broderick J, Olson S, Shaw S, Spilker J, Meyer B, Emr M, Warren M, Marler J. NIHSS training and certification using a new digital video disk is reliable. *Stroke*. 2005;36:2446–2449.
18. Kasner SE. Clinical interpretation and use of stroke scales. *Lancet Neurol*. 2006;5:603–612.