CrossMark

ORIGINAL ARTICLE

Web-Based Assessment of Outcomes After Subarachnoid and Intracerebral Hemorrhage: A New Patient Centered Option for Outcomes Assessment

**Andrew M. Naidech · Jennifer L. Beaumont · Michael Berman ·
Eric Liotta · Matthew B. Maas · Shyam Prabhakaran · Konrad Kording ·
Jane Holl · David Cella**

Abstract

Background Clinical outcomes are typically assessed by trained staff. We tested the hypothesis that outcomes reported by the patient or a caregiver on the web would be correlated with a validated interview.

Methods We assessed surviving patients with intracerebral and subarachnoid hemorrhage at 1-, 3-, and 12-month follow-up with a validated interview for the modified Rankin Scale (mRS, a validated ordinal scale from 0, no symptoms to 5, severe disability). Health-related quality of life (HRQoL) was assessed on the web with NIH Patient-Reported Outcomes Measurement Information System (PROMIS) and Neuro-QOL using computer adaptive testing by the patient, proxy reporting by a caregiver, or proxy entry by study staff.

Results A coincident mRS and HRQoL assessment was available for 149 (71 %) of 209 patients at one, three, or 12 months. There were 89 assessments with proxy entry by study staff, 89 by the patient on the web, and 58 with proxy report by a caregiver on the web. PROMIS physical function assessments were completed in median of 4 questions, and *T* scores were associated with the mRS ($P < 0.001$), regardless of respondent. Mean *T* scores in every category of the mRS were different from every other category ($P \leq 0.003$ for all). Results were similar for Neuro-QOL mobility.

Conclusions Web-based HRQoL assessment with NIH PROMIS and Neuro-QOL is feasible and correlated with a validated interview for the mRS. *T* scores distinguished between individual categories of the mRS, detecting modest differences in physical function and mobility HRQoL that are difficult to detect with the mRS. PROMIS and Neuro-QOL provide powerful and sensitive outcomes for potentially large cohorts.

Keywords Outcomes research ·
Intracerebral hemorrhage · Subarachnoid hemorrhage ·
Quality of life

A. M. Naidech (✉) · M. Berman · E. Liotta ·
M. B. Maas · S. Prabhakaran
Department of Neurology, Northwestern University, 710 N Lake
Shore Drive 11th floor, Chicago 60611, USA
e-mail: a-naidech@northwestern.edu

A. M. Naidech · J. L. Beaumont · D. Cella
Departent of Medical Social Sciences, Northwestern University,
Chicago, USA

A. M. Naidech · J. Holl
Center for Healthcare Studies, Institute for Public Health and
Medicine, Northwestern University, Chicago, USA

K. Kording
Rehabilitation Institute of Chicago, Chicago, IL, USA

Introduction

A validated assessment of outcomes after discharge is crucial to understand the impact of acute treatment. The current standard for patients with cerebrovascular disease is a global ordinal scale of functional outcomes, typically the modified Rankin Scale (mRS) at one [1] or, more typically, 3 months. It is widely used in studies of acute ischemic stroke [2], intracerebral hemorrhage (ICH) [3], and subarachnoid hemorrhage (SAH). The mRS has many advantages [4], including a high inter-rater reliability [5] and validated questionnaires for in-person [6, 7] or telephone interview [8] of the patient or a caregiver. Other ordinal scores, such as the Glasgow Outcome Score, are similar.

There are drawbacks, however, to the use of a validated interview for outcome assessment. A dedicated interview for each patient can be time intensive, especially in large study populations, increasing costs. Additionally, the mRS may not be patient centered. For example, "good outcome" in studies of patients with SAH or ICH often means walking independently, but better cognitive function (e.g., the ability to manage one's own finances) may be more important to an individual. Thus, there is substantial room for improvement in outcomes assessment.

There are opportunities to improve the analysis of outcomes as well. Ordinal outcomes are usually dichotomized for logistic regression, although this reduces statistical power and cannot detect differences within the categories of "good" and "poor" outcome [9]. For example, an improvement from bed bound to limited mobility in a wheelchair would not be captured, as both are considered "poor outcome." Ordinal regression to detect smaller improvements is statistically more powerful than logistic regression, but requires the data meet assumptions (specifically, proportional odds) that are difficult to guarantee in advance. A widely adopted continuous numeric outcomes score would have important statistical advantages [10].

The internet has been useful in collecting clinical outcomes data [11] by reducing demands on researchers and patients. If patients and caregivers could reliably report outcomes electronically, then outcomes ascertainment might be performed with reduced effort, burden, and cost. Whether or not these patient-reported or caregiver-reported outcomes are valid compared to the mRS is not well described, and this has held back the acceptance of web-assessed outcomes.

The NIH has invested substantially to develop the Patient-Reported Outcomes Measurement Information System (PROMIS), composed of over two dozen validated instruments to assess mobility, applied cognition, and other specific domains. Neuro-QOL [12] is a set of instruments particularly germane to neurological disorders, and substantially overlaps with PROMIS. Both provide valid and reliable tools for health-related quality of life (HRQoL) assessment. We tested the hypothesis that patients or proxies could independently report HRQoL on the web with NIH PROMIS and Neuro-QOL, and that HRQoL T scores would be associated with the mRS as assessed by a validated interview.

Methods

Patients

We prospectively enrolled consecutive patients from January 2011 through January 2014. All patients had a diagnosis of spontaneous intracerebral hemorrhage (ICH) or subarachnoid hemorrhage (SAH) confirmed by a board-certified neurologist and head computed tomography (CT). Patients with trauma, hemorrhagic conversion of ischemic stroke, or structural lesions (e.g., tumor, arteriovenous malformation, vessel dissection) were excluded. We recorded the medical history, severity of injury including the NIH Stroke Scale (NIHSS), and demographic information.

Procedure

We approached patients or a legally authorized representative during the index hospitalization and asked for written consent to track identifiers and obtain outcomes, a preferred telephone number and email addresses. When PROMIS and Neuro-QOL became available for research use in January 2011, we attempted to obtain the mRS and HRQoL at 1 and 3 months. In late May 2011, we began to attempt follow-up at 12 months.

mRS Assessment

The mRS is a validated scale from 0 (no symptoms) to 6 (death), with a score of 4 indicating dependence. A single interviewer (MB) obtained the mRS by validated interview [6]. For patients no longer in the hospital, the mRS was assessed by telephone interview, a commonly used method validated by others [4, 8, 13–15].

HRQoL Assessment

Coincident with the mRS assessment we sent an email with a link to complete the HRQoL assessment, the usual method. The respondent was asked to identify himself/herself as the patient, a caregiver, study staff, etc. If the respondent was not the patient, he/she was given specific instructions to answer what the patient would answer. Respondents could also answer HRQoL questions over the telephone with study staff (MB) performing proxy entry if web-based reporting was not available or inconvenient. Only one HRQoL assessment was counted at each time point, i.e., if a patient responded we did not seek a proxy report from a caregiver. We did not seek HRQoL assessment if the patient was known to have died. Study staff only recorded what the person said and could not draw inferences.

We assessed HRQoL with PROMIS physical function, a bank of 124 items, and Neuro-QOL lower extremity function (mobility), a bank of 19 items. Both were assessed with computer adaptive tests, where the response determines subsequent questions, so that only a subset of questions was required for each respondent. For example, a respondent who indicates no difficulty walking may be asked about difficulty running, while a respondent who reports difficulty walking may be asked

Table 1 Characteristics of 114 outcomes assessed at 1 month

	Patient reported on web	Reported on web by caregiver	Entered on web by study staff	P for difference
N	42	27	45	
Age	47.5 ± 13.1	65 ± 13.6	60.9 ± 16	<0.001
NIH stroke scale on admission	0 [0–2]	6 [1–20]	3 [1–14]	<0.001
mRS by interview	1 [0–1]	4 [3–5]	3 [1–4]	<0.001
Ethnicity				0.4
White	34 (81)	20 (74)	29 (64)	
Black	3 (7)	6 (22)	10 (22)	
Asian	5 (12)	1 (4)	5 (12)	
Women	20 (47)	13 (48)	18 (40)	0.4
History of coronary artery disease	1 (2)	5 (19)	4 (9)	0.07
History of atrial fibrillation	1 (2)	0	2 (4)	0.5
History of hypertension	15 (36)	17 (63)	24 (53)	0.07
History of diabetes	6 (14)	4 (14)	6 (13)	0.9

Patients who reported their own outcomes on the web were younger, had a lower (better) NIH Stroke Scale on admission, and had less functional disability as evidenced by a lower modified Rankin Scale (mRS). There were similar differences for the 59 patients assessed at one year

Data are N (%), mean ± SD, or median [Q1–Q3] as appropriate

about difficulty standing. Results are expressed in T scores, centered on the general US population at 50 ± 10. Further information regarding the computer adaptive testing algorithm and instruments is available at www.assessmentcenter.net, www.nihpromis.org, and www.neuroqol.org. Both are also available in validated short forms of 6–10 questions and the entire item bank, although we did not use them here. In this report, we focused on mobility and physical function because they are most directly measured by the mRS. We did not mandate a defined criteria for "computer literacy" to take part.

Statistical Analysis

T Score is a continuous number normalized to the US general population at 50 ± 10. We tested the hypothesis that, at any given time point (1-, 3-, or 12-month follow-up), HRQoL T scores would be associated with the mRS at that time. We used analysis of variance to test this hypothesis. Multiple comparisons, such as comparing the T scores for patients with mRS 0 versus mRS 1, mRS 0 versus 2, mRS 0 versus 3, etc. were corrected for multiple comparisons with the Least Significant Differences technique. Non-normally distributed data (e.g., NIHSS scores between groups) were compared with the Kruskal–Wallis statistic. Calculations were made with standard statistical software (IBM SPSS v. 22, Armonk, NY). A statistician from Neuro-QOL and the PROMIS Statistical Center who was not involved in the acquisition of data (JLB) directed and reviewed the statistical analysis.

Results

We identified coincident HRQoL T scores and mRS in 149 (71 %) of 209 patients. We excluded patients for whom there was a mRS but no HRQoL data (34, including 11 who died), patients for whom there were HRQoL data but no mRS (23), and patients for whom no follow-up information could be ascertained (3).

These 209 patients had 236 total HRQoL assessments; 89 (38 %) assessments were proxy entry by study staff; 89 (38 %) by the patient on the web; and 58 (24 %) by caregiver report from a caregiver on the web, 55 of whom were family members. Thus, most HRQoL reporting was independent of study staff.

There were 114 (48 %) assessments at 1 month, 63 (27 %) assessments at 3 months, and 59 (25 %) assessments at 12 months. HRQoL assessments at 1 month were completed promptly: 82 (72 %) on the same day; 21 (18 %) within 3 days; and 10 (9 %) within a week. Data were similar for 1 year.

There was no association between the respondent and the time of assessment ($P = 0.3$). For example, the patient responded in 42 of 114 (37 %) assessments at 1 month, 18 of 63 (28 %) assessments at 3 months, and 28 of 59 (47 %) assessments at 12 months.

HRQoL assessments required modest effort to complete. The PROMIS physical function computer adaptive test was completed in a median [Q1–Q3] of 4 [4, 5] questions. There was no association between the number of questions

administered and the respondent ($P = 0.4$). There were similar results for Neuro-QOL mobility.

Not surprisingly, the respondent was associated with outcomes. Patients who reported their own HRQoL outcomes at 1 month had a lower (better) NIHSS at admission, were younger, and had a lower (better) mRS (Table 1). There were similar associations at 1 year. No patient who reported his/her own HRQoL was classified as mRS = 5 (very severe disability).

HRQoL T scores were linearly associated with the mRS (Fig. 1, $P < 0.001$). The associations were similar for both PROMIS general physical function and Neuro-QOL mobility HRQoL. After correction for multiple comparisons, PROMIS physical function and Neuro-QOL mobility T scores in each category of the mRS were significantly different from scores in every other category of the mRS (Table 2, $P \leq 0.003$ for all). The regression line for T score of physical function was 53.1–6.7* (mRS at Assessment); the addition of the respondent did not add to this model ($P = 0.4$). In other words, each increase in the mRS by one point was associated with a decrease in physical function of 6.7 points, or 0.67 SD.

Results were similar when each time of assessment (1, 2, or 12 months) or the respondent was considered separately.
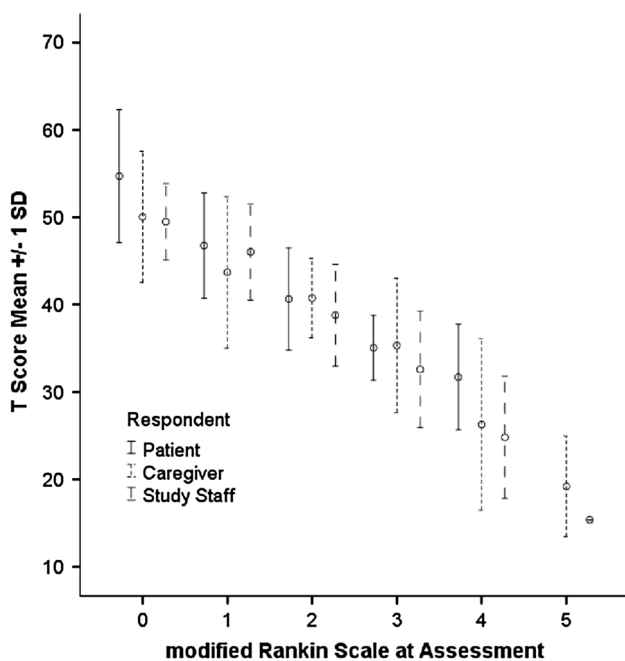

Fig. 1 Mean T scores ±1 standard error for NIH PROMIS physical function T scores (y axis) plotted against the modified Rankin Scale (mRS, scored from 0, no symptoms to 5, severe disability) at the time of assessment. T scores were associated with the mRS ($P < 0.001$), showing higher (worse) mRS associated with lower physical function T scores. Data reported on the web by the patient (*solid*), caregiver as proxy (*stippled*) and study staff as proxy (*dash*) are shown, showing the similar relationship between mRS and T scores regardless of the respondent. No self-identified patient with mRS of 5 reported his/her own HRQoL, so there are no patient-reported data for mRS 5. Data for Neuro-QOL mobility T scores are similar

Table 2 T scores for PROMIS physical function and Neuro-QOL mobility health-related quality of life, stratified by the modified Rankin Scale (mRS) at the time of assessment

Modified Rankin Scale	N (%)	PROMIS physical function T score	Neuro-QOL mobility T score
0—No symptoms	50 (21)	52.7 ± 7.1	57.5 ± 6.3
1—No disability	50 (21)	46.1 ± 6.1	52.1 ± 7.9
2—Mild disability	35 (15)	39.9 ± 5.6	47.1 ± 7.6
3—Moderate disability, independent	41 (17)	33.9 ± 6.5	38.8 ± 7.4
4—Moderately severe disability, dependent	39 (17)	26.1 ± 8.5	27.4 ± 9.4
5—Very severe disability	21 (9)	17.6 ± 4.7	16.6 ± 3.0

T scores were different between categories of the mRS overall ($P < 0.001$) and after correction for multiple comparisons ($P \leq 0.003$ for all), showing that higher (worse) mRS scores by validated interview were reflected in T scores obtained from web-based assessment. Results were similar when the analysis was restricted to a single time of assessment (1, 3, or 12 months) or a single respondent (patient, proxy report by caregiver, proxy entry by study staff)

Data are N (%) or mean ± SD

Discussion

These data demonstrate that HRQoL responses for survivors of ICH and SAH could be reported on the web, and that the resulting T scores were associated with the mRS by validated interview. T scores varied linearly with the mRS, underscoring the progressive measure of disability. These data lend further support to web-based patient-reported outcomes as a valid option for outcomes assessment for patients with cerebrovascular disease.

T scores from NIH PROMIS and Neuro-QOL have distinct advantages. The patient or a family member reported HRQoL data in 147 (62 %) assessments, allowing us to obtain these data with reduced effort compared to a validated interview. The modest number of questions needed to calculate a T score from computer adaptive testing indicates a low burden on the respondent.

HRQoL T scores within each category of the mRS were similar regardless of the respondent. Neuro-QOL has been validated for proxy report as part of its development and we did not intend to repeat these validation studies, rather, we assessed its correlation with the interview for the mRS. These results suggest PROMIS physical function is also valid for proxy entry, probably due to the objective nature of the questions.

Patient respondents were generally independent, while caregiver respondents usually acted as proxy for a disabled

patient. While this may lead to some bias, this is an unavoidable component of assessing outcomes of disabling conditions. The interview for the mRS and any other functional assessment will have a similar potential bias. As expected, patients with very severe disability (mRS = 5) required someone to report outcomes on their behalf, and their HRQoL scores were dismal.

Study staff did not provide feedback to patients or caregivers regarding their level of ability during the assessment of the mRS that could potentially bias HRQoL reporting. Coaching would likely have been ineffective in any event, since the PROMIS general physical function instrument contains questions ranging in difficulty from sitting on the edge of a bed to running ten miles. The computer adaptive testing algorithm chooses the questions based on the respondent's previous answers, making "cheating" difficult even if there were a secondary gain to achieving a score. The modest number of questions administered by the computer adaptive testing algorithm suggests that estimation was straightforward. PROMIS physical function and Neuro-QOL mobility HRQoL scores were similar, although not exactly the same.

While the mRS will continue to serve as a reference standard, T scores for HRQoL have several statistical advantages vis-à-vis the mRS by dichotomous or ordinal regression, as one can use techniques for continuous variables [10]. It is plausible that there are interventions that are not statistically powerful enough to increase the odds of "good outcome" on the mRS but are powerful enough to significantly improve HRQoL T scores. NIH PROMIS and Neuro-QOL may therefore improve the statistical power to show an effect of various future interventions through more sensitive outcomes assessment. Some interventions that have previously not improved the odds of "good outcome" might improve HRQoL T scores; as an example, aggressive blood pressure reduction improved HRQoL scores, but not "good outcome" on the mRS in patients with acute ICH [16]. These advantages make HRQoL an attractive primary endpoint for translational and clinical research.

We focused on physical scores in this study, but have previously noted that Neuro-QOL assessments highlight other domains of HRQoL that are unlikely to be adequately assessed by the mRS, such as applied cognition (now referred to as "cognitive function" within Neuro-QOL) [17]. How well the mRS captures these other domains of HRQoL is an important topic for future research, but beyond the scope of this report.

We sent individual scripted emails, although a bulk email could also be sent to a cohort of patients. While this study assessed outcomes for a moderate sized registry, a web-based assessment could be used to assess HRQoL for a much larger cohort of patients such as in a large epidemiologic study, patients with common insurance coverage,

and so on. The NIH Toolbox and PROMIS Assessment Center specifically have a method for presenting a consent form and collecting identifiers on the web, and many Institutional Review Boards (including ours) permit online consent, although written consent was obtained in this study. Online security at Assessment Center is robust, and similar to that for online banking, bill payment, and health insurance. While digital security is a concern, one cannot steal another's identity by knowing their physical function T Score and personalized identifying information such as a social security number is not required to participate. There is also less risk of undesired consequences of HRQoL reporting since having pre-existing disability is no longer a reason to deny insurance coverage.

This technique for outcomes assessment has broad implications across cerebrovascular disease and clinical neurosciences. Although we assessed patients with ICH and SAH in this cohort, one of us (SP) directs the outcome assessment of patients with acute ischemic stroke, and study staff has consistently obtained HRQoL by proxy entry over the telephone because it was not known if independent report would be reliable when we started. These data may lead us to reconsider this time-intensive approach, and suggest data from any respondent would be equivalent.

There are limitations to web-based assessment. Internet access may be unreliable, patients and caregivers may not feel comfortable or have the skills to complete an assessment on the web. Not all patients may have easy access to a computer and Assessment Center is not yet optimized for smartphones, although this is a potential future project. Outcomes in these cases were generally assessed by proxy entry by study staff reading the questions over the telephone. NIH PROMIS physical function and Neuro-QOL mobility are available as short forms on paper, but these do not take advantage of computer adaptive testing and require later data entry. HRQoL assessment requires active participation, as opposed to a completely automated system for outcomes assessment such as retrieval of discharge disposition from hospital records, activity monitors, posts on social media, and so on. Using social media postings to assess patients is likely to be biased in favor of patients with better outcomes. This study concentrated on physical function and mobility, but other domains of HRQoL, such as cognitive function, may also be important to patients and caregivers. The mRS interview was not audio recorded for quality assurance or the identity of the person reporting the mRS, although previous investigations have developed validated questionnaires to minimize any potential bias. Errors in identifying the respondent would not have changed the association between T scores and the mRS by validated interview, since the association between the mRS and HRQoL T Score did not vary with the respondent.

In sum, this study found that patients and caregivers could independently report HRQoL in the domains of physical functional and mobility, and that these were correlated with a validated interview for the mRS. Web-based HRQoL assessment is a method to record outcomes on a large scale, leading to more patient centered and statistically powerful clinical research.

Conflict of interest The authors are unaware of any conflicts of interest.

References

1. Ovbiagele B, Lyden PD, Saver JL, Collaborators FTV. Disability status at 1 month is a reliable proxy for final ischemic stroke outcome. Neurology. 2010;75:688–92.
2. Hacke W, Kaste M, Bluhmki E, Brozman M, Davalos A, Guidetti D, Larrue V, Lees KR, Medeghri Z, Machnig T, Schneider D, von Kummer R, Wahlgren N, Toni D. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med. 2008;359:1317–29.
3. Mayer SA, Brun NC, Broderick J, Davis S, Diringer MN, Skolnick BE, Steiner T. Efficacy and safety of recombinant activated factor vii for acute intracerebral hemorrhage. New Engl J Med. 2008;358:2127–37.
4. Weimar C, Kurth T, Kraywinkel K, Wagner M, Busse O, Haberl RL, Diener H-C. Assessment of functioning and disability after ischemic stroke. Stroke. 2002;33:2053–9.
5. Banks JL, Marotta CA. Outcomes validity and reliability of the modified rankin scale: implications for stroke clinical trials: a literature review and synthesis. Stroke. 2007;38:1091–6.
6. Saver JL, Filip B, Hamilton S, Yanes A, Craig S, Cho M, Conwit R, Starkman S, Investigators f t F-M, Coordinators. Improving the reliability of stroke disability grading in clinical trials and clinical practice: the rankin focused assessment (rfa). Stroke. 2010;41:992–5.
7. Wilson JTL, Hareendran A, Grant M, Baird T, Schulz UGR, Muir KW, Bone I. Improving the assess of outcomes in stroke. use of a structured interview to assign grade on the modified rankin scale. Stroke. 2002;33:2243–6.
8. Bruno A, Akinwuntan AE, Lin C, Close B, Davis K, Baute V, Aryal T, Brooks D, Hess DC, Switzer JA, Nichols FT. Simplified modified rankin scale questionnaire: reproducibility over the telephone and validation with quality of life. Stroke. 2011;42:2276–9.
9. Bath PM, Gray LJ, Collier T, Pocock S, Carpenter J. Can we improve the statistical analysis of stroke trials? Statistical reanalysis of functional outcomes in stroke trials. Stroke. 2007;38:1911–5.
10. Bath PM, Lees KR, Schellinger PD, Altman H, Bland M, Hogg C, Howard G, Saver JL. Statistical analysis of the primary outcome in acute stroke trials. Stroke. 2012;43:1171–8.
11. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. Nat Biotechnol. 2011;29:411–4.
12. Cella D, Lai JS, Nowinski CJ, Victorson D, Peterman A, Miller D, Bethoux F, Heinemann A, Rubin S, Cavazos JE, Reder AT, Sufit R, Simuni T, Holmes GL, Siderowf A, Wojna V, Bode R, McKinney N, Podrabsky T, Wortman K, Choi S, Gershon R, Rothrock N, Moy C. Neuro-qol: brief measures of health-related quality of life for clinical research in neurology. Neurology. 2012;78:1860–7.
13. Chen KF, Colantuoni E, Siddiqi F, Dinglas VD, Sepulveda KA, Fan E, Pronovost PJ, Needham DM. Repeated attempts using different strategies are important for timely contact with study participants. J Clin Epidemiol. 2011;64:1144–51.
14. Dennis M, Mead G, Doubal F, Graham C. Determining the modified rankin score after stroke by postal and telephone questionnaires. Stroke. 2011;. doi:10.1161/STROKEAHA.111.639708.
15. Mayer SA, Kreiter KT, Copeland D, Bernardini GL, Bates JE, Peery S, Claassen J, Du YE, Connolly ES Jr. Global and domain-specific cognitive impairment and outcome after subarachnoid hemorrhage. Neurology. 2002;59:1750–8.
16. Anderson CS, Heeley E, Huang Y, Wang J, Stapf C, Delcourt C, Lindley R, Robinson T, Lavados P, Neal B, Hata J, Arima H, Parsons M, Li Y, Heritier S, Li Q, Woodward M, Simes RJ, Davis SM, Chalmers J. Rapid blood-pressure lowering in patients with acute intracerebral hemorrhage. N Engl J Med. 2013;368:2355–65.
17. Naidech AM, Beaumont JL, Rosenberg NF, Maas MB, Kosteva AR, Ault ML, Cella D, Ely EW. Intracerebral hemorrhage and delirium symptoms: length of stay, function and quality of life in a 114-patient cohort. Am J Respir Crit Care Med. 2013;188:1331–7.