J Neurol (2006) 253 : 788–793
DOI 10.1007/s00415-006-0119-x

ORIGINAL COMMUNICATION

Development and validation of a prognostic model to predict recovery following intracerebral hemorrhage

Christian Weimar
Michael Roth
Vera Willig
Panagiotis Kostopoulos
Jens Benemann
Hans-Christoph Diener
for the German Stroke Study Collaboration

Received: 2 August 2005
Received in revised form: 5 December 2005
Accepted: 15 December 2005
Published online: 13 March 2006

Ch. Weimar, MD (✉) · J. Benemann, PhD · H.-Ch. Diener, MD
Department of Neurology
University of Duisburg-Essen
Hufelandstr. 55
45122 Essen, Germany
Tel.: +49-201/723-2369
Fax: +49-201/723-5753
E-Mail: stroke.med@uni-essen.de

M. Roth, MD · P. Kostopoulos, MD
Department of Neurology
Universitätskliniken des Saarlandes
66421 Homburg-Saar, Germany

V. Willig, MD
Department of Neurology
University of Jena
Philosophenweg 3
07743 Jena, Germany

Collaborators of the German Stroke Study Collaboration
Neurology departments (responsible for study investigators):
Krankenanstalten Gilead Bielefeld (C. Hagemeister, MD), Rheinische Kliniken Bonn (C. Kley, MD), University of Saarland (P. Kostopoulos, MD), University of Jena (V. Willig, MD), University of Magdeburg (M. Goertler, MD), Klinikum Minden (J. Glahn, MD), Städtisches Krankenhaus München Harlaching (K. Aulich, MD), University of Rostock (A. Kloth, MD), Bürgerhospital Stuttgart (T. Mieck, MD), University of Ulm (M. Riepe, MD), University of Essen (V. Zegarac, MD)

Abstract *Context* While several models have been developed to predict mortality following intracerebral hemorrhage (ICH), the functional outcome and its predictors in surviving patients have been poorly investigated so far. *Objectives* To identify predictors and validate a prognostic model for independent functional outcome in patients with acute ICH. *Design* An inception cohort was assessed on the National Institutes of Health Stroke Scale (NIH-SS) at admission and followed-up after 100 days. *Setting* 11 neurological departments with an acute stroke unit. *Patients* 207 consecutive patients who were neither comatose nor intubated at admission within 6 hours after ICH and with complete follow-up. *Results* After 100 days, 40 patients (19.3 %) had died, 78 (37.7 %) had regained functional independence (Barthel Index ≥ 95) and 89 (43 %) had survived but not recovered. In these patients, age and the NIH-SS total score were identified as independent predictors for functional independence after 100 days. With the predefined cut-off value, the prognosis of 79.8 % of all patients could be predicted accurately upon validation in an independent data set of 173 non-comatose patients with acute ICH. *Conclusion* Our study provides a validated prognostic model for prediction of complete recovery following ICH which could be very useful for the design of clinical studies.

Key words intracerebral hemorrhage · outcome · prediction · prognosis · NIH-SS

Background

Pooled data from community based studies indicate a 1-month mortality after primary intracerebral hemorrhage of 40–45 % [8]. Because of this high fatality and the relatively low incidence only few patients are available for follow-up assessment of functional outcome in survivors. Moreover, early neurological examination and cerebral imaging are rarely available in all patients included in these studies and are often delayed for several days by which time secondary worsening may have already occurred. Hospital-based studies therefore can provide valuable information on initial stroke severity and outcome for larger patient numbers. Identification of prognostic factors for functional outcome in patients not comatose at admission would be especially useful for the design of randomized clinical trials to estimate

JON 2119

sample size and define inclusion criteria. In addition, validated prognostic models enable adjustment for case mix to compare outcome and quality of care in different institutions [5]. Finally, this information could also be helpful as an individual prognostic indication both to relatives and physicians. We therefore focused our study on functional outcome in patients admitted within 6 hours after onset of ICH. In a second step, the model was validated in an independent data set from a previous study [23].

Methods

The 11 neurological departments listed in the appendix participated in this study. Enrollment of patients started on 1 July 2000, and was terminated on 15 March 2002. Details on data collection and management have been previously described [22]. All patients were treated according to best current knowledge in clinical routine. Operative evacuation of hematoma or extraventricular drainage was performed in 12 (4%) patients. Imaging studies were performed to diagnose patients with primary hemorrhage and identify possible etiologies. Greatest horizontal diameter of hemorrhage was measured on site by either a neuroradiologist or the treating neurologist on axial slices of the first cerebral imaging. Patients or their next of kin were informed about study participation and written consent was obtained to forward personal data to the coordinating center. The study was approved by the Ethics Committee of the University of Essen and aspects of data safety were approved by the responsible data protection state representative. The admitting physician's bedside prediction for outcome after 100 days was assessed in one of the following categories: death, severe dependence (Barthel Index < 70), moderate dependence (Barthel Index 70–90), and functional independence (Barthel Index ≥ 95). Only predictions made within the first 24 hours after admission were considered for analysis.

Severity of stroke was assessed on the National Institutes of Health Stroke Scale (NIH-SS) at admission as well as 48–72 hours after admission [16]. Investigators were experienced in the use of the NIH-SS by video training and other clinical studies. As suggested by an internationally agreed definition, we investigated worsening in any one of the following NIH-SS items: level of consciousness, gaze, arm and leg motor function and speech [4]. A central follow up blinded to baseline variables was performed via telephone interview by the coordinating center or by the treating hospital itself, if the patient did not give consent that personal data be forwarded. Patient outcome was assessed on the Barthel Index (BI) 100 (median 102) days after the event or by confirmation of death within 120 days after initial stroke. If no follow-up information could be obtained from either the patient, relatives or the treating physicians, local death registries were screened.

Only patients admitted within 6 hours after onset of ICH (N = 327) and with a prior Rankin grade ≤ 2 (N = 303) were included in this study to ensure that patients were functionally independent to a certain degree. Of 43 patients intubated or comatose at admission, 19 patients died within 72h after admission, overall 28 patients died within 100 days and only 2 patients had completely recovered with a Barthel Index ≥ 95 after 100 days. Because of this highly adverse prognosis, we included only patients who were neither intubated nor comatose at admission to ensure a realistic chance of independent functional outcome. Of the remaining 260 patients, 53 (20.4%) could not be reached for a complete follow-up. Patients lost to follow-up were not significantly different from patients with complete follow-up information (Table 1). The flow chart of patient inclusion is depicted in Fig. 1.

Statistical analysis was performed with the program package SPSS Version 10.0. If a single variable was not available for all patients only valid cases were reported. If more than 1% of cases were missing, the

Table 1 Comorbid conditions and clinical findings in patients not intubated nor comatose at admission with functional dependence or death, functional independence (Barthel Index, BI ≥95), and missing follow-up at 100 days after ICH

Variable (missing cases in patients with follow-up)	Died or dependent (N = 129)	BI ≥95 (N = 78)	lost to follow-up (N = 53)
Mean age/SD	69.1/9.6*	60.5/12.0	64.7/15.2
Women, %	44.2	35.9	37.7
Prior stroke, %	18.6	17.9	15.1
Prior coronary heart disease, %	19.4*	9.0	9.4
Arterial hypertension, %	77.5	75.6	75.5
Diabetes mellitus, %	20.9	11.5	13.2
Coagulation disorder, %	3.9	2.6	3.8
Lobar hemorrhage, %	34.1	37.2	32.1
Lenticulostriate hemorrhage, %	34.9	24.4	34.0
Thalamic hemorrhage, %	17.8	17.9	18.9
Putaminal hemorrhage, %	8.5	9.0	11.3
Cerebellar hemorrhage, %	3.9	6.4	1.9
Brainstem hemorrhage, %	5.4	7.7	3.8
Axial diameter > 4 cm, % (1)	35.2*	11.5	24.5
Ventricular bleeding or hydrocephalus, %	21.7*	3.8	20.8
Alert at admission, %	47.3*	87.2	58.5
Mean NIH-SS at admission/Median	12.6/12*	6.4/6	10.7/11
Neurological worsening within 72 h, % (5)	32.0*	6.5	14.3
Fatal bedside prognosis, % (14)	8.2*	0	0
Favorable bedside prognosis, % (14)	7.4*	31.0	20.8

* significant with $p < 0.05$ compared to patients with Barthel Index ≥95

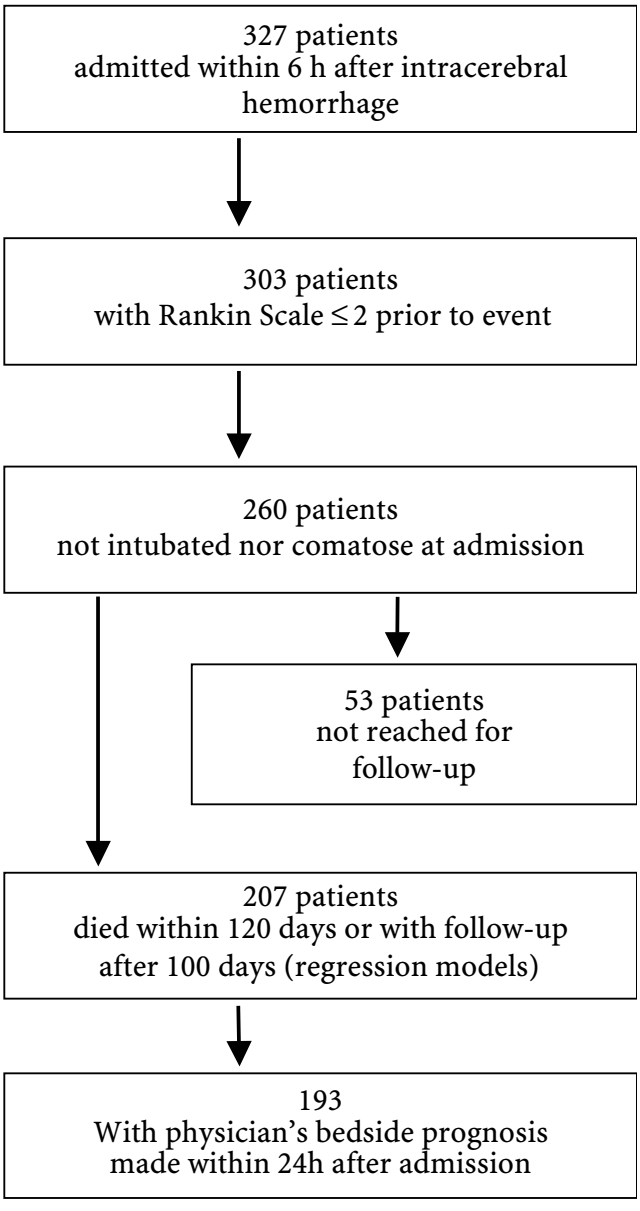

Fig. 1 Patient inclusion chart

number of missing cases was additionally provided. To test for statistical differences, Student's t-test was used to compare age, the χ^2-test to compare categorical variables and the Mann-Whitney U-test to compare stroke severity on the NIH-SS. We developed a logistic regression model for prediction of independent versus dependent functional outcome (BI < 95) or death 100 days after admission. We included all significant variables from the univariate analyses which could be assessed within the first hours after admission. Thus, 7 variables (age, sex, prior coronary heart disease, initial NIH-SS total score, initial NIH-SS level of consciousness, axial diameter of bleeding, ventricular bleeding and/or hydrocephalus) were fitted into the logistic regression model via forward and backward stepwise selection. The number of events per variable was above 10, as recommended in order to avoid over fitting the regression model. Nevertheless, variables were retained only if their resulting p-value was < 0.005 [1, 24]. From models with all variables that resulted from any of the selection procedures, any variable with $p > 0.005$ was eliminated stepwise. To the remaining set of variables, every previously eliminated variable was again added and kept in the model if it fulfilled the same criteria. Finally, all two-way interactions of the resulting variables were investigated and kept if $p < 0.005$. In the final model, regression coefficients with standard errors, odds ratio and 95 % confidence intervals are reported. The threshold for classification using the logistic distribution function was set so that the predicted proportion of events was equal to the observed. We assessed the discrimination of the regression models by calculating the area under the receiver operating characteristic (ROC) curve, which is a plot of sensitivity of predictions against 1-specificity of predictions. An area under the ROC curve (AUC) of 0.5 indicates no discrimination (i. e., the line follows the 45° diagonal), and an area of 1.0 (i. e., the line includes the entire area within the horizontal and vertical axes) indicates perfect discrimination.

The model was validated in an independent cohort of patients with ICH which have been previously described [23]. In short, 586 patients with ICH were documented consecutively in 30 hospitals during a one year period between 1998–1999. For model validation we excluded patients admitted more than 6h after ictus (N = 291), prior Rankin Scale > 2 (N = 30), comatose level of consciousness at admission (N = 39) and with unknown functional status after 3 months (N = 53). The predictive model with the previously calibrated threshold was validated in the remaining 173 patients with acute ICH.

Results

260 patients met the baseline inclusion criteria and were included in this study. Within 72 hours after admission, 15 patients (5.8 %) had died, 19 patients (7.3 %) were newly intubated, 18 patients (6.9 %) had worsening of consciousness and 9 patients were not reassessed within 72 hours after admission. 52 patients with worsening of key neurological functions or death after 72 hours were significantly less often alert at admission, had a higher NIH-SS total score at admission, greater diameter of hemorrhage and more often ventricular bleeding (all significant). After 100 days, 40 patients (19.3 %) had died, 78 (37.7 %) had regained functional independence (Barthel Index ≥ 95), 89 (43 %) had survived but not recovered. In 53 patients a complete follow-up was missing. Baseline characteristics of these patients were not significantly different from all patients with complete follow-up. Compared with patients with complete recovery, patients with unfavorable outcome were significantly older, more frequently had a history of prior coronary heart disease, greater diameter of bleeding, more often ventricular bleeding and/or hydrocephalus, a higher NIH-SS total score at admission and were less often alert at admission (Table 1). The relationship between age and unfavorable outcome is shown in Fig. 2.

A logistic regression model in 207 patients who were neither comatose nor intubated at admission and with complete follow-up after 100 days identified the NIH-SS total score at admission and age as independent predictors for functional independence (Table 2). The resulting model had an optimal cut-off at 0.42 and an area under the ROC curve of 0.861 (Fig. 3). While this model correctly predicted 81.6 % of all patients, the treating physi-

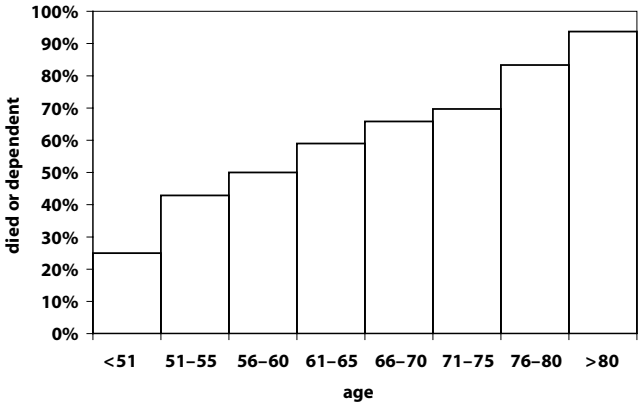

Fig. 2 relationship between age and unfavorable outcome (Barthel Index < 95 or death) after 100 days of 207 patients in the development cohort

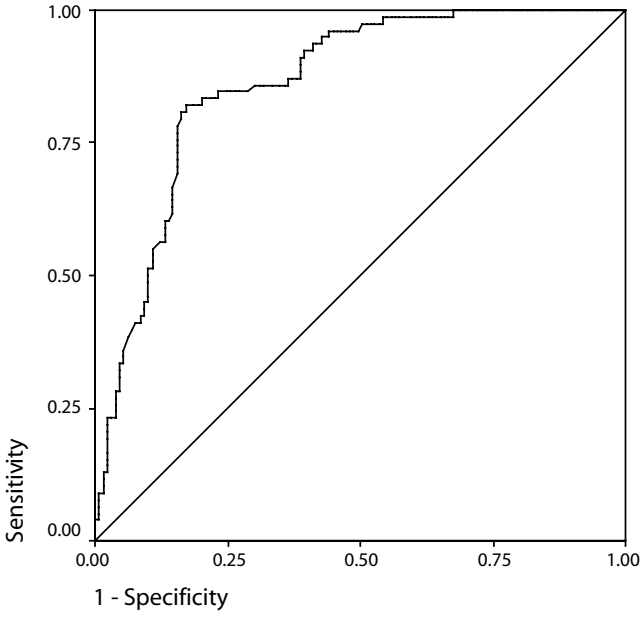

Fig. 3 Receiver operating curve of the regression model for functional independence (Barthel Index ≥95) after 100 days in 207 patients not intubated or comatose at admission (model development)

cians within the first 6 hours after admission correctly predicted only 31 % of patients who regained functional independence and 92.6 % of patients who died or did not regain functional independence (overall correct prognosis in 69.9 %).

Upon validation in an independent patient cohort from another hospital-based study, the resulting area under the ROC curve was 0.876 (Fig. 4). Upon use of the pre-specified threshold, a correct prediction was obtained of 91.9 % of patients who recovered and 63.5 % of patients who did not fully recover. The calibrated percentage of correctly classified patients is presented in Table 3.

Discussion

Exact outcome estimations from validated prognostic models can help to define an optimized study population for both obtaining sufficient statistical power and limiting study duration [21, 25]. We focused on complete recovery and mortality 100 days after an acute ischemic stroke as endpoints of primary interest and selected the Barthel Index as the most widely used measure of functional independence. Our study had a well defined cohort and predominantly central follow-up with external validation in an independent patient cohort. By excluding patients with no realistic chance of reaching functional restitution, our model reflects a typical study population for acute ICH trials. While not all patients could be reached for follow-up, no significant differences in the initial physicians' prognosis as well as other baseline variables could be found between patients lost and with complete follow-up. In addition, death registries were screened before any patient was considered lost to follow-up. Nevertheless, patients with serious disability tend to be harder to reach for follow-up and therefore we cannot completely exclude a follow-up bias. However, this should not have influenced the validity of our results which can therefore be considered representative for patients with ICH treated in German departments of neurology. We were unable to include patients from neurosurgery departments which are more likely to receive patients needing operative evacuation or extraventricular drainage. Because the STICH trial did not show any significant benefit of early surgery versus initial conservative treatment [18], it is unlikely that surgical treatment in a greater percentage of patients would have had a major influence on our results. In addition, neurosurgical patients may not constitute a suitable study population for medical intervention trials.

Table 2 Regression model for functional independence (Barthel Index ≥95) after 100 days in 207 patients not intubated nor comatose at admission

Variable	β	S. E.	Odds ratio	95 % CI	p-value
Age (1 year increase)	–0.089	0.020	0.915	0.880–0.951	< 0.001
NIH-SS total score	–0.244	0.041	0.783	0.722–0.849	< 0.001
Intercept	7.543	1.444			

β regression coefficient; *S. E.* standard error; *CI* confidence interval

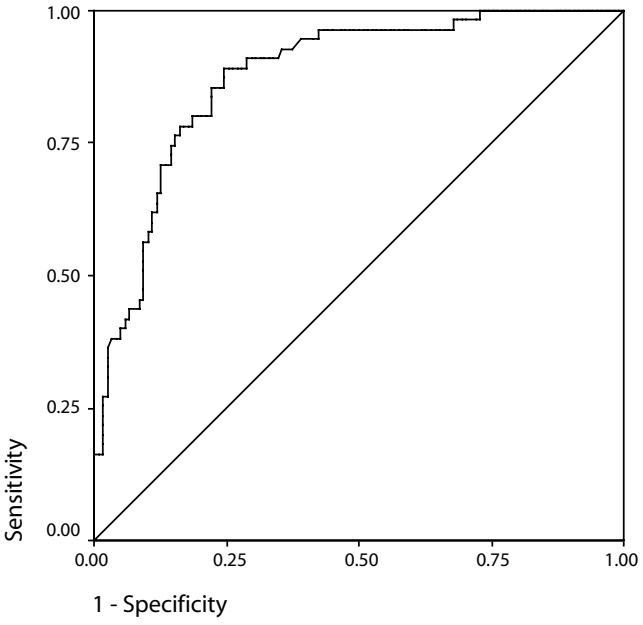

Fig. 4 Receiver operating curve of the regression model for functional independence (Barthel Index ≥95) after 100 days in 173 patients with ICH (validation cohort)

Numerous predictors for outcome following ICH have been proposed. Among the most consistently identified are extent of hemorrhage, initial Glasgow Coma Scale (GCS), presence of ventricular bleeding, age, location of hemorrhage, sex, global clinical impression, midline shift, hydrocephalus, history of arterial hypertension, and glucose level during the first days (review by Ariesen et al. [2]). Numerous ICH scoring systems have been proposed, only some of which were designed to predict functional outcome [6, 7, 11, 15, 19, 20]. Only one score has been externally validated with contradictory results concerning the accuracy of favorable outcome following ICH [7, 9, 13, 14]. In contrast to previous findings, only age and the initial NIH-SS could be identified as independent predictors of independent functional outcome. As shown in a previous study, the NIH-SS had a higher predictive value for three month outcome than any other indicator of stroke severity [7]. This scale is widely accepted in clinical studies of acute stroke and enables a more discriminative assessment of neurological deficits than the GCS. In addition, the GCS is biased towards aphasic patients who score considerably worse than non-aphasic patients. The NIH-SS therefore could be more useful for clinical studies than a score that focuses on severely affected patients but has only minor prognostic value in the rest. Interestingly, the imaging parameters included in model development were of lower significance than age and initial NIH-SS as a clinical measure of stroke severity and were not retained in the final model. While hematoma volume and location of hematoma have been previously suggested as independent predictors of good outcome, the respective models did not exclude patients with highly adverse prognosis (i. e. intubated or comatose at admission) and used less rigorous levels of significance [6, 7, 11, 15, 19, 20]. Although imaging variables would probably have been retained in a larger development cohort, their inclusion would have added little diagnostic accuracy but instead would have increased complexity and rendered the model less stable. Likewise, the prediction could probably be improved by including variables that are assessed at a later stage of treatment or rehabilitation but the practical value would be limited as the prediction cannot be given as early. We therefore decided to base our model on variables which could be obtained within the first hours after admission following ICH.

Physicians in our study tended to be overly pessimistic regarding the outcome of patients with ICH. Other studies have suggested that do-not-resuscitate (DNR) orders may lead to self-fulfilling prophecies [3, 12]. Although we had no information on DNR orders or withdrawal of support, this seems unlikely to have affected our population after exclusion of comatose or intubated patients. Few studies have investigated the impact of early neurological worsening upon outcome after ICH [10, 17]. Inclusion of a variable to account for neurological worsening after 72 hours yielded an additional independent predictor of adverse functional outcome when added to the identified models (data not shown).

In conclusion, neurologists tend to be overly pessimistic regarding the chance of functional restitution after ICH. Our validated model for recovery after 100 days therefore could assist in providing quantitative prognosis to patients and their families. The results could also provide valuable information for the design of future clinical trials by excluding patients with a high

Table 3 Classification of 173 patients in the validation cohort using the previously calibrated threshold of 0.42

Observed frequency (%)	Predicted frequency (%)		
	Barthel Index ≥95	Adverse outcome	Total
BI ≥95	91 (52.6)	8 (4.6)	99 (57.2)
Adverse outcome	27 (15.6)	47 (27.2)	74 (42.8)
Total	118 (68.2)	55 (31.8)	173

chance of spontaneous recovery who are thus unlikely to show a treatment effect.

Acknowledgements This study was sponsored by the German Ministry of Education and Research (BMBF) as part of the Competence Net Stroke and the Deutsche Forschungsgemeinschaft (DI 327/8–1). The funding sources had no involvement in the study. We thank Peter Dommes, PhD for central data collection and management.

References

1. Altman DG, Lyman GH (1998) Methodological challenges in the evaluation of prognostic factors in breast cancer. Breast Cancer Res Treat 52:289–303
2. Ariesen MJ, Algra A, van der Worp HB, Rinkel GJ (2005) Applicability and relevance of models that predict short term outcome after intracerebral haemorrhage. J Neurol Neurosurg Psychiatry 76:839–844
3. Becker KJ, Baxter AB, Cohen WA, Bybee HM, Tirschwell DL, Newell DW, Winn HR, Longstreth WT Jr (2001) Withdrawal of support in intracerebral hemorrhage may lead to self-fulfilling prophecies. Neurology 56:766–772
4. Birschel P, Ellul J, Barer D (2004) Progressing stroke: Towards an internationally agreed definition. Cerebrovasc Dis 17:242–252
5. Cadilhac DA, Ibrahim J, Pearce DC, Ogden KJ, McNeill J, Davis SM, Donnan GA (2004) Multicenter comparison of processes of care between Stroke Units and conventional care wards in Australia. Stroke 35: 1035–1040
6. Castellanos M, Leira R, Tejada J, Gil-Peralta A, Davalos A, Castillo J (2005) Predictors of good outcome in medium to large spontaneous supratentorial intracerebral haemorrhages. J Neurol Neurosurg Psychiatry 76: 691–695
7. Cheung RT, Zou LY (2003) Use of the original, modified, or new intracerebral hemorrhage score to predict mortality and morbidity after intracerebral hemorrhage. Stroke 34:1717–1722
8. Dennis MS (2003) Outcome after brain haemorrhage. Cerebrovasc Dis 16 (Suppl 1):9–13
9. Fernandes H, Gregson BA, Siddique MS, Mendelow AD (2002) Testing the ICH score. Stroke 33:1455–1456; author reply 1455–1456
10. Flemming KD, Wijdicks EF, St Louis EK, Li H (1999) Predicting deterioration in patients with lobar haemorrhages. J Neurol Neurosurg Psychiatry 66:600–605
11. Hallevy C, Ifergane G, Kordysh E, Herishanu Y (2002) Spontaneous supratentorial intracerebral hemorrhage. Criteria for short-term functional outcome prediction. J Neurol 249:1704–1709
12. Hemphill JC 3rd, Newman J, Zhao S, Johnston SC (2004) Hospital usage of early do-not-resuscitate orders and outcome after intracerebral hemorrhage. Stroke 35:1130–1134
13. Hemphill JC, Bonovich DC, Besmertis L, Manley GT, Johnston SC, Tuhrim S (2001) The ICH Score: A simple, reliable grading scale for intracerebral hemorrhage. Stroke 32:891–897
14. Jamora RD, Kishi-Generao EM Jr, Bitanga ES, Gan RN, Apaga NE, San Jose MC (2003) The ICH score: predicting mortality and functional outcome in an Asian population. Stroke 34:6–7; author reply 6–7
15. Lisk DR, Pasteur W, Rhoades H, Putnam RD, Grotta JC (1994) Early presentation of hemispheric intracerebral hemorrhage: prediction of outcome and guidelines for treatment allocation. Neurology 44:133–139
16. Lyden P, Brott T, Tilley B, Welch KM, Mascha EJ, Levine S, Haley EC, Grotta J, Marler J (1994) Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. Stroke 25:2220–2226
17. Mayer SA, Sacco RL, Shi T, Mohr JP (1994) Neurologic deterioration in noncomatose patients with supratentorial intracerebral hemorrhage. Neurology 44:1379–1384
18. Mendelow AD, Gregson BA, Fernandes HM, Murray GD, Teasdale GM, Hope DT, Karimi A, Shaw MD, Barer DH (2005) 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 365:387–397
19. Portenoy RK, Lipton RB, Berger AR, Lesser ML, Lantos G (1987) Intracerebral haemorrhage: a model for the prediction of outcome. J Neurol Neurosurg Psychiatry 50:976–979
20. Schwarz S, Hafner K, Aschoff A, Schwab S (2000) Incidence and prognostic significance of fever following intracerebral hemorrhage. Neurology 54:354–361
21. Weimar C, Ho TW, Katsarava Z, Diener HC (2006) Improving Patient Selection for Clinical Acute Stroke Trials. Cerebrovasc Dis 21:386–392
22. Weimar C, Koenig IR, Kraywinkel K, Ziegler A, Diener HC (2004) Age and National Institutes of Health Stroke Scale Score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. Stroke 35:158–162
23. Weimar C, Weber C, Wagner M, Busse O, Haberl RL, Lauterbach KW, Diener HC (2003) Management patterns and health care use after intracerebral hemorrhage. a cost-of-illness study from a societal perspective in Germany. Cerebrovasc Dis 15:29–36
24. Weimar C, Ziegler A, Koenig I, Diener H (2002) Predicting functional outcome and survival after acute ischemic stroke. J Neurol 249:888–895
25. Weir CJ, Kaste M, Lees KR (2004) Targeting neuroprotection clinical trials to ischemic stroke patients with potential to benefit from therapy. Stroke 35:2111–2116