

## PAPER

# Development and validation of the Essen Intracerebral Haemorrhage Score

C Weimar, J Benemann, H-C Diener, for the German Stroke Study Collaboration\*



See Editorial Commentary, p 571

*J Neural Neurosurg Psychiatry* 2006;77:601–605. doi: 10.1136/jnnp.2005.081117

See end of article for authors' affiliations

Correspondence to:  
Dr Christian Weimar,  
Department of Neurology,  
University of Duisburg-  
Essen, Hufelandstr 55,  
45122 Essen, Germany;  
stroke.med@uni-essen.de

Received  
26 September 2005  
In revised form  
5 December 2005  
Accepted 5 December 2005  
Published Online First  
14 December 2005

**Background:** Spontaneous intracerebral haemorrhage (ICH) accounts for the highest in-hospital mortality of all stroke types. Nevertheless, outcome is favourable in about 30% of patients. Only one model for the prediction of favourable outcome has been validated so far.

**Objective:** To describe the development and validation of the Essen ICH score.

**Methods:** Inception cohorts were assessed on the National Institutes of Health stroke scale (NIH-SS) on admission and after follow up of 100 days. On the basis of previously validated clinical variables, a simple clinical score was developed to predict mortality and complete recovery (Barthel index after 100 days  $\geq 95$ ) in 340 patients with acute ICH. Subscores for age (<60=0; 60–69=1; 70–79=2;  $\geq 80=3$ ), NIH-SS level of consciousness (alert=0; drowsy=1; stuporose=2; comatose=3), and NIH-SS total score (0–5=0; 6–10=1; 11–15=2; 16–20=3;  $>20$  or coma=4) were combined into a prognostic scale with  $<3$  predicting complete recovery and  $>7$  predicting death. The score was subsequently validated in an external cohort of 371 patients. **Results:** The Essen ICH score showed a high prognostic accuracy for complete recovery and death in both the development and validation cohort. For prediction of complete recovery on the Barthel index after 100 days, the Essen ICH score was superior to the physicians' prognosis and to two previous prognostic scores developed for a slightly modified outcome.

**Conclusions:** The Essen ICH score provides an easy to use scale for outcome prediction following ICH. Its high positive predictive values for adverse outcomes and easy applicability render it useful for individual prognostic indications or the design of clinical studies. In contrast, physicians tended to predict outcome too pessimistically.

In view of a rather poor short term prognosis of patients with intracerebral haemorrhage (ICH), most prognostic models so far have focused on the prediction of mortality.<sup>1</sup> For clinical stroke trials, however, an increase in favourable outcome is a more meaningful end point than a decrease in mortality. In addition, prognostic models for recovery can be helpful in individual cases both for relatives and physicians. Finally, validated prognostic models based on easily available clinical variables enable adjustment for case mix to compare outcome and quality of care in different institutions.<sup>2</sup>

Of the five previously developed scores to predict favourable versus unfavourable outcome or death following ICH,<sup>3–8</sup> only the ICH score<sup>8</sup> has been validated so far. We have recently developed and validated a logistic regression model to predict complete recovery (Barthel index  $\geq 95$ ) after 100 days in patients with acute ICH, which identified age and the National Institutes of Health stroke scale (NIH-SS) total score on admission as the sole independent predictors. For model development we had included all significant variables from univariate analysis such as history, initial stroke severity, and cerebral imaging which could be routinely assessed within the first hours after admission.<sup>9</sup> To further improve the applicability of our model, we sought to translate the regression model into a general and easily applicable predictive scale, and subsequently validate the resulting 11 point Essen ICH score in an independent dataset. In a next step, we compared its performance with the only other validated score for predicting a good outcome, the ICH score, along with a recently proposed improved version, the modified ICH score (table 1).<sup>7, 8</sup>

## METHODS

### Model development

We selected the Barthel index (BI) as the most widely used measure of functional outcome.<sup>10</sup> This scale evaluates

individual abilities in feeding, dressing, mobility (walking on a level surface and ascending/descending stairs), and personal hygiene (grooming, toileting, bathing, and control of bodily functions). It thus adequately reflects functional consequences for daily activities that are immediately important to the patient. To identify patients with complete recovery as advocated for clinical trials, a cut off BI value of  $\geq 95$  v  $<95$  was used.

As a first step, we developed a logistic regression model for predicting complete recovery on the BI after 100 days.<sup>9</sup> For developing this model we considered all variables on previous history, stroke severity, and imaging information which could be assessed routinely within the first hours after admission. Only age and the NIH-SS were retained as independent predictors of complete recovery after 100 days in non-comatose patients admitted within six hours after ICH.

As a second step, we designed a simple and more general predictive score, including patients admitted up to 24 hours after ictus as well as comatose patients, using the data bank of the German Stroke Foundation which has been described previously.<sup>11</sup> In short, 586 patients with ICH were documented consecutively in 30 hospitals during a one year period between 1998 and 1999 with a predominantly central follow up telephone interview after three months. Severity of stroke was assessed on the NIH-SS at admission.<sup>12</sup> Investigators were experienced in the use of the NIH-SS through video

**Abbreviations:** AUC, area under the curve; BI, Barthel index; ICH, intracerebral haemorrhage; NIH-SS, National Institutes of Health stroke scale; ROC, receiver operating characteristic; STICH, International Surgical Trial in Intracerebral Haemorrhage

\* Collaborators of the German Stroke Study Collaboration are listed at the end of the article

**Table 1** Prognostic scores for complete recovery and death following intracerebral haemorrhage

Predictors	Essen ICH score	ICH score <sup>8</sup>	Modified ICH score <sup>7</sup>
Age	<60=0 60–69=1 70–79=2 ≥80=3	<80=0 ≥80=1	<80=0 ≥80=1
NIH-SS	0–5=0 6–10=1 11–15=2 16–20=3 >20 or coma = 4		0–10=0 11–20=1 21–40=2
NIH-SS level of consciousness	Alert=0 Drowsy=1 Stupor=2 Coma=3		
Glasgow coma scale (GCS)		3–4=2 5–12=1 13–15=0	
Haemorrhage volume (cm <sup>3</sup> )		≥30=1 <30=0	≥30=1 <30=0
Infratentorial origin		Yes=1 No=0	Yes=1 No=0
Intraventricular haemorrhage		Yes=1 No=0	Yes=1 No=0
Maximum score	10	6	6
Death predicted	>7	≥3	≥3
Good outcome predicted	<3	<3	<3
Definition of good outcome	Barthel index ≥95	Rankin scale ≤2	Rankin scale ≤2
Follow up (days)	100	30	30

ICH, intracerebral haemorrhage; NIH-SS, National Institutes of Health stroke scale.

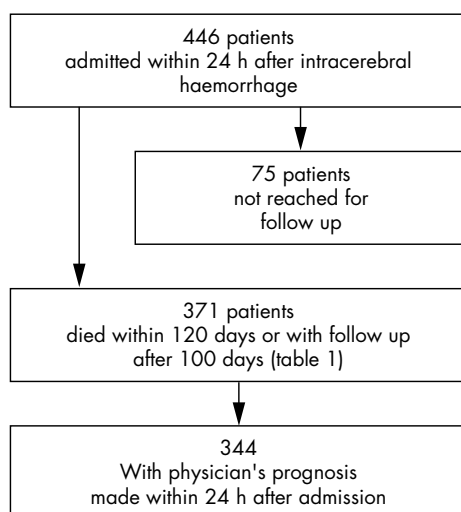
training and other clinical studies. We excluded patients admitted more than 24 hours after ictus ( $n = 120$ ) and those without follow up information ( $n = 126$ ). Patients could not be reached for follow up mainly because of limited funding for the central follow up, as well as lack of staff in the participating centres to pursue local inquiries. Patients lost to follow up were not significantly different with regard to age, sex, or initial stroke severity on the NIH-SS from those with complete follow up. Of the remaining 340 patients with complete follow up, 126 (37.1%) had died and 89 (26.2%) had recovered completely (BI  $\geq 95$ ). In this population we

designed a prognostic scale based on the two previously identified independent predictors age and NIH-SS on admission.

To account for the highly adverse prognosis of intubated or comatose patients, level of consciousness was included as a third variable despite overlapping with the NIH-SS total score. Multiple models with different subscores for each of the three predictive variables were run with different cut off values for mortality and complete recovery. The best fit in the development dataset was obtained by adding up subscores for age (<60 = 0; 60–69 = 1; 70–79 = 2;  $\geq 80 = 3$ ), NIH-SS level of consciousness (alert = 0; drowsy = 1; stupor = 2; comatose = 3), and NIH-SS total score (0–5 = 0; 6–10 = 1; 11–15 = 2; 16–20 = 3; >20 or coma = 4) (table 1). On a range from 0 to 10, an Essen ICH score of >7 best predicted mortality and a score of <3 best predicted complete recovery.

### Model validation

Subsequently, the Essen ICH score was validated in an independent hospital based consecutive cohort which was assessed in all hospitals participating in a prospective validation study of predictive models for ischaemic stroke.<sup>13</sup> The 11 neurological departments listed at the end of the paper contributed data to this study and included more than 90% of all patients with ICH admitted within 24 hours after ictus. Enrolment of patients started on 1 July 2000, and was terminated on 15 March 2002. All patients received routine clinical treatment according to best current knowledge. Imaging studies were done to diagnose primary haemorrhage. The greatest horizontal diameter of haemorrhage was measured on site by either a neuroradiologist or the treating neurologist on axial slices of the first cerebral imaging study. Patients or their next of kin were informed about study participation and written consent was obtained to forward



**Figure 1** Patient inclusion chart (validation cohort).

**Table 2** characteristics of patients with ICH in the development and validation cohort

Variable	Development (n = 340)	Validation (n = 371)
Age (years) (mean (SD))	65.9 (12.6)	67.4 (12)
Women (%)	43.8	43.9
Rankin scale >1 before ICH (%)	13.4	15.1
History of stroke (%)	16.0	19.4
History of arterial hypertension (%)	66.8	73.3
History of diabetes mellitus (%)	15.6	17.5
Alert on admission (%)	46.2	54.2*
Comatose on admission (%)	18.8	15.6
NIH-SS on admission (mean (median))	15.3 (13)	13.2 (10)*
Axial diameter ≥4 cm (%)	NA	30.6
Lobar localisation (%)	NA	36.4
Infratentorial localisation (%)	NA	11.6
Ventricular bleeding (%)	NA	39.5
Death during hospital stay (%)	27.4	20.2*
Death within 120 d after admission (%)	37.1	29.1*
BI ≥95 after 100 days (%)	26.2	32.9

\*Significant at p<0.05.  
 BI, Barthel index; ICH, intracerebral haemorrhage; NA, not assessed; NIH-SS, National Institutes of Health stroke scale.

personal data to the coordinating centre. 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.

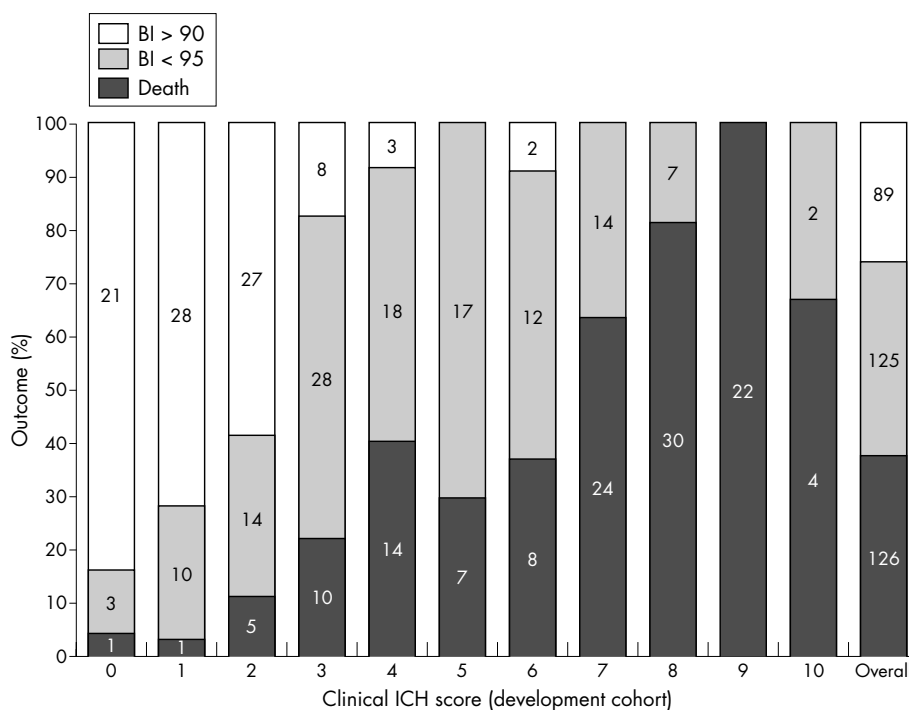
The admitting physician’s prediction of outcome after 100 days was placed in one of the following categories: death, severe dependence (BI <70), moderate dependence (BI 70–90), and complete recovery (BI ≥95). Only predictions made within the first 24 hours after admission were considered for analysis.

A central follow up blinded to baseline variables was made by telephone interview by the coordinating centre or by the treating hospital itself, if the patient did not consent for personal data to be forwarded. Patient outcome was assessed

on the BI at 100 days (median 103.5) or by confirmation of death within 120 days after the event. Death registries were screened if no follow up information could be obtained. Patients lost to follow up (16.8%) were not significantly different with respect to age, sex, initial NIH-SS, or the initial physician’s prognosis compared with patients with complete follow up.

The flow chart of patient inclusion is shown in fig 1. If the NIH-SS score was not available in patients who were intubated at admission, these patients were considered to be comatose for the purpose of calculating the Essen ICH score.

We finally compared the predictive accuracy of our Essen ICH score, the physicians’ prediction, the ICH score, and the modified ICH score (table 1).<sup>7, 8</sup> In order to apply the ICH score to our validation dataset, we aggregated stroke severity on the NIH-SS into the three categories of the ICH score, which is based on the Glasgow coma scale. For calculation of the modified ICH score, haemorrhage volume was assumed to be >30 cm<sup>3</sup> if the maximum diameter on axial slices was ≥4 cm. The ICH score and modified ICH score were developed for prediction of good outcome, defined as ≤2 on the modified Rankin scale. Because no information on the Rankin scale at follow up was available in the validation dataset, we instead estimated the predictive accuracy of these scales for complete recovery, defined as BI ≥95. As this end point is harder to achieve than a score ≤2 on the modified Rankin scale,<sup>14</sup> the defined cut off value of the ICH score and modified ICH score may not be optimal, resulting in a wider difference between the sensitivity and specificity than described in the original publication.<sup>7</sup> To compare these scores with our Essen ICH score and the physicians’ prediction we therefore assessed the discrimination of the various prognostic scores 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 (that is, the line follows the 45° diagonal), and an area of 1.0 (that is, the line includes the entire area within the horizontal and vertical axes) indicates perfect discrimination.



**Figure 2** 100-Day outcome on the Essen ICH score in 340 patients (model development). BI, Barthel index; ICH, intracerebral haemorrhage.

**Table 3** Performance of the Essen ICH score, physicians' prediction, ICH score,<sup>8</sup> and modified ICH score<sup>7</sup> for prediction of death and complete recovery (Barthel index  $\geq 95$ ) in the validation dataset (n = 371)

	Essen ICH score		Physicians' prediction (n = 344)		ICH score		Modified ICH score	
	Death	Complete recovery	Death	Complete recovery	Death	Complete recovery	Death	Complete recovery
Sensitivity (%)	43.9	73.8	41.2	37.8	58.5	64.2	64.2	64.2
Specificity (%)	97.7	84.2	98.8	95.3	93.1	85.8	85.8	85.8
PPV (%)	88.7	69.8	93.3	79.2	77.5	64.8	64.8	64.8
NPV (%)	81.0	86.7	80.1	76.5	84.7	85.5	85.5	85.5
Area under ROC curve (95% CI)	0.831 (0.784 to 0.878)	0.877 (0.842 to 0.912)	0.812 (0.762 to 0.861)	0.831 (0.786 to 0.876)	0.831 (0.783 to 0.878)	0.776 (0.729 to 0.824)	0.835 (0.787 to 0.882)	0.813 (0.769 to 0.858)

\*Models designed to predict a modified Rankin scale of  $\leq 2$  but in this table used to predict a Barthel index of  $\geq 95$ . CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic.

**Statistical analysis**

Statistical analysis was carried out with the program package SPSS version 10.0 (Cary, North Carolina, USA). To test for statistical differences, Student's *t* test was used to compare age, the  $\chi^2$  test to compare categorical variables, and the Mann-Whitney U test to compare stroke severity on the NIH-SS.

**RESULTS**

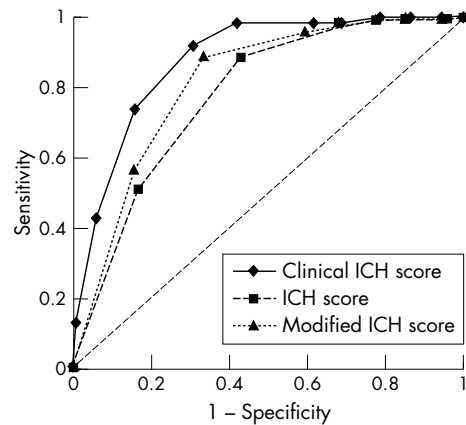
Patient characteristics of the development and validation cohort are shown in table 2.

Patients in the development cohort had a greater initial stroke severity and a higher in-hospital and 120 day mortality than patients in the validation cohort ( $p < 0.05$ ). In the development cohort, a predefined Essen ICH score of  $>7$  best predicted mortality (sensitivity 44.4%, specificity 95.8%, positive predictive value (PPV) 86.2%, negative predictive value (NPV) 74.5%) and a score of  $<3$  best predicted complete recovery (sensitivity 85.4%, specificity 86.5%, PPV 69.1; NPV 94.3%) (fig 2). The ROC curve yielded an AUC of 0.851 for death and 0.913 for complete recovery. When applying the Essen ICH score to the validation cohort, the ROC curve was 0.831 for death and 0.877 for complete recovery. With the predefined cut off values, the sensitivity for death was 43.9% (PPV 88.7%) and for complete recovery 73.8% (NPV 86.7%) (table 3).

In contrast, the physicians' prediction had a sensitivity of 41.2% (PPV 93.3%) for death and 37.8% (NPV 76.5) for complete recovery. Both the classical ICH score and the modified ICH score had an almost identical AUC in the ROC statistics for death but performed worse in predicting complete recovery on the BI after 100 days (table 3, fig 3).

**DISCUSSION**

We developed and validated a simple clinical score for predicting complete recovery and death following ICH. Our comparably large cohorts for model development and validation were well defined and had a predominantly central follow up. Owing to the setting in hospitals of different levels of care, patients in the development cohort had a greater stroke severity, as well as a worse outcome, compared with the validation cohort from neurology departments with acute stroke units. Although not all patients could be reached for follow up, the outcomes in the validation cohort should be representative of patients with



**Figure 3** Receiver operating characteristic curves of the Essen ICH score, ICH score, and modified ICH score for prediction of complete recovery (Barthel index  $\geq 95$ ) after 100 days in the validation dataset (n = 371). ICH score and modified ICH score were models designed to predict a modified Rankin scale of  $\leq 2$  but in this graph were used to predict a Barthel index of  $\geq 95$ .



ICH treated in German neurology departments with acute stroke units. We were unable to include patients from neurosurgery or emergency departments, who are likely to have a different prognosis. Validation of our score in these patients is therefore still required. While only a small minority of our patients received surgical evacuation of haematoma, surgical treatment in a greater percentage of patients would have been unlikely to have influenced prognosis, as the STICH trial did not show any significant benefit of early surgery versus initial conservative treatment.<sup>15</sup> Because the NIH-SS is difficult to assess in comatose patients, we assigned the worst subscore to patients with either total NIH-SS >20 or comatose level of consciousness which facilitates the use of the Essen ICH score. Patients intubated on admission had a similarly adverse prognosis, with 79% mortality and only 4% having a favourable outcome after 100 days; in the validation these were considered to be comatose for calculating the Essen ICH score. Nevertheless, intubated patients who are awake or only drowsy on admission (which is rather hypothetical during the first 24 hours after ictus) should be assessed on the NIH-SS according to the official instructions.

For prediction of good outcome, only the ICH score had previously been validated in an independent dataset.<sup>7,8</sup> Without three dimensional measurements of haemorrhage extension for calculating the ICH score and the modified ICH score, we had to rely on volume estimates from the axial diameter. While this may partly explain the low sensitivity of these scores for death in the validation cohort compared with the original publication, a lower threshold for this variable does not have any importance for predicting a good outcome.<sup>7</sup> Because we did not have information on the 30 day modified Rankin scale we were unable to make a direct comparison of the sensitivity and specificity of the ICH score and modified ICH score with our Essen ICH score. However, the ROC statistics, which include different thresholds on each scale, showed a superior discrimination of our Essen ICH score for predicting complete recovery after 100 days. Although the use of the previous scores in this study was not as the models were designed or validated, the comparison offers a valuable estimate of how the Essen ICH score performs. Moreover, our score was superior to the physicians' prediction given within the first 24 hours. Nevertheless, it remains to be determined how the Essen ICH score performs in predicting the 30 day modified Rankin scale compared with the previous models. Although sensitivity with the predefined cut off for death was only moderate and lower than the sensitivity of the ICH score and the modified ICH score, the high specificity and positive predictive value of this prediction may be important for individual prognoses for relatives or for making therapeutic decisions. In contrast, the even higher positive predictive value of the physicians' prediction of death may hint at a self fulfilling prophecy.<sup>16-18</sup>

## Conclusion

The rather pessimistic physicians' prediction for a favourable outcome supports the use of a prognostic scale for treatment decisions. The Essen ICH score could also provide valuable prognostic indications for patients and relatives. Finally, it could improve the design of future clinical trials by either defining prognosis adjusted end points<sup>15</sup> or excluding patients with a high chance of spontaneous recovery, who are thus unlikely to show any measurable treatment effect.

## COLLABORATORS OF THE GERMAN STROKE STUDY COLLABORATION

Neurology departments (responsible study investigators) are as follows: Krankenanstalten Gilead Bielefeld (C Hagemester), Rheinische Kliniken Bonn (C Kley), University of Saarland (P Kostopoulos), University of Jena (V Willig), University of Magdeburg (M Goertler), Klinikum Minden (J Glahn), Städtisches Krankenhaus München Harlaching (K Aulich), University of Rostock (A Kloth), Bürgerhospital Stuttgart (T Mieck), University of Ulm (M Riepe), University of Essen (V Zegarac).

## 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 Klaus Kraywinkel MD MSc and Peter Dommies PhD for central data collection and management.

## Authors' affiliations

C Weimar, J Benemann, H-C Diener, Department of Neurology, University of Duisburg-Essen, Essen, Germany

Competing interests: none declared

## REFERENCES

- 1 Ariesen MJ, Algra A, van der Worp HB, *et al.* Applicability and relevance of models that predict short term outcome after intracerebral haemorrhage. *J Neural Neurosurg Psychiatry* 2005;**76**:839-44.
- 2 Cadilhac DA, Ibrahim J, Pearce DC, *et al.* Multicenter comparison of processes of care between Stroke Units and conventional care wards in Australia. *Stroke* 2004;**35**:1035-40.
- 3 Hallevy C, Ifergane G, Kordysh E, *et al.* Spontaneous supratentorial intracerebral hemorrhage. Criteria for short-term functional outcome prediction. *J Neurol* 2002;**249**:1704-9.
- 4 Schwarz S, Hafner K, Aschoff A, *et al.* Incidence and prognostic significance of fever following intracerebral hemorrhage. *Neurology* 2000;**54**:354-61.
- 5 Lisk DR, Pasteur W, Rhoades H, *et al.* Early presentation of hemispheric intracerebral hemorrhage: prediction of outcome and guidelines for treatment allocation. *Neurology* 1994;**44**:133-9.
- 6 Portenoy RK, Lipton RB, Berger AR, *et al.* Intracerebral haemorrhage: a model for the prediction of outcome. *J Neural Neurosurg Psychiatry* 1987;**50**:976-9.
- 7 Cheung RT, Zou LY. Use of the original, modified, or new intracerebral hemorrhage score to predict mortality and morbidity after intracerebral hemorrhage. *Stroke* 2003;**34**:1717-22.
- 8 Hemphill JC, Bonovich DC, Besmertis L, *et al.* The ICH Score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke* 2001;**32**:891-7.
- 9 Weimar C, Roth M, Willig V, *et al.* Development and validation of a prognostic model to predict recovery following intracerebral hemorrhage. *J Neural*, (in press).
- 10 Mahony FI, Barthel DW. Functional evaluation: the Barthel Index. *Md Med J* 1965;**14**:61-5.
- 11 Weimar C, Weber C, Wagner M, *et al.* Management patterns and health care use after intracerebral hemorrhage. a cost-of-illness study from a societal perspective in Germany. *Cerebrovasc Dis* 2003;**15**:29-36.
- 12 Lyden P, Brott T, Tilley B, *et al.* Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. *Stroke* 1994;**25**:2220-6.
- 13 German Stroke Study Collaboration. Predicting outcome after acute ischemic stroke: an external validation of prognostic models. *Neurology* 2004;**62**:581-5.
- 14 Uyttenboogaart M, Stewart RE, Vroomen PC, *et al.* Optimizing cutoff scores for the Barthel index and the modified Rankin scale for defining outcome in acute stroke trials. *Stroke* 2005;**36**:1984-7.
- 15 Mendelow AD, Gregson BA, Fernandes HM, *et al.* 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**:387-97.
- 16 Becker KJ, Baxter AB, Cohen WA, *et al.* Withdrawal of support in intracerebral hemorrhage may lead to self-fulfilling prophecies. *Neurology* 2001;**56**:766-72.
- 17 Hemphill JC, Newman J, Zhao S, *et al.* Hospital usage of early do-not-resuscitate orders and outcome after intracerebral hemorrhage. *Stroke* 2004;**35**:1130-4.
- 18 Zurasky JA, Aiyagari V, Zazulia AR, *et al.* Early mortality following spontaneous intracerebral hemorrhage. *Neurology* 2005;**64**:725-7.