Predicting Long-Term Outcome After Acute Ischemic Stroke

A Simple Index Works in Patients From Controlled Clinical Trials

Inke R. König, PhD; Andreas Ziegler, PhD; Erich Bluhmki, PhD; Werner Hacke, MD;
Philip M.W. Bath, MD; Ralph L. Sacco, MD, MS; Hans C. Diener, MD; Christian Weimar, MD;
on behalf of the Virtual International Stroke Trials Archive (VISTA) Investigators*

Background and Purpose—An early and reliable prognosis for recovery in stroke patients is important for initiation of individual treatment and for informing patients and relatives. We recently developed and validated models for predicting survival and functional independence within 3 months after acute stroke, based on age and the National Institutes of Health Stroke Scale score assessed within 6 hours after stroke. Herein we demonstrate the applicability of our models in an independent sample of patients from controlled clinical trials.

Methods—The prognostic models were used to predict survival and functional recovery in 5419 patients from the Virtual International Stroke Trials Archive (VISTA). Furthermore, we tried to improve the accuracy by adapting intercepts and estimating new model parameters.

Results—The original models were able to correctly classify 70.4% (survival) and 72.9% (functional recovery) of patients. Because the prediction was slightly pessimistic for patients in the controlled trials, adapting the intercept improved the accuracy to 74.8% (survival) and 74.0% (functional recovery). Novel estimation of parameters, however, yielded no relevant further improvement.

Conclusions—For acute ischemic stroke patients included in controlled trials, our easy-to-apply prognostic models based on age and National Institutes of Health Stroke Scale score correctly predicted survival and functional recovery after 3 months. Furthermore, a simple adaptation helps to adjust for a different prognosis and is recommended if a large data set is available. **(*Stroke*. 2008;39:000-000.)**

Key Words: stroke ■ outcome ■ National Institutes of Health Stroke Scale ■ prediction ■ clinical trials

The importance of an early and reliable prognosis for recovery in patients with acute stroke is undisputed.[1,2] In addition to clinical reasons, such as information for patients and family as well as adapting treatment and rehabilitation options, inclusion of prognostic information in controlled clinical trials helps to define individual clinical end points, to select suitable patients, and to reduce required sample sizes.[3–5]

To be useful and applicable to clinical practice, a prognostic model needs to be validated and easy to implement; ie, it should contain only a few variables that are readily available for all patients.[6] A systematic review that included studies until 1997 showed that the methodology of most reported prognostic models for stroke recovery was poor, and none of the models was recommended for clinical practice or research.[7] Since then, additional prognostic models have been developed. Of these, the validated models by Baird et al[8] and Johnston et al[9–11] that predicted recovery within 3 months relied on imaging variables, which may not be available for all patients and in all settings. In contrast, the models by Counsell et al[12] and our own group[13] included a few simple clinical variables and were also subsequently validated.[14–16]

Because in these models prognostic variables were assessed within a delay of 4 (median) and 2 to 3 days after stroke, respectively, timely prediction for initiating acute treatment was precluded. To allow for an almost immediate prognosis based on a few simple variables, we recently developed and externally validated models for predicting survival and functional independence within 3 months.[17] These models are based on age and neurologic impairment as measured on the National Institutes of Health Stroke Scale (NIHSS), assessed within 6 hours after stroke onset. In the external validation sample, ≈75% of patients were classified correctly for functional independence and >85% with regard to survival.

In some respects, patients in our previous training and validation samples were highly comparable: all patients were admitted consecutively to German neurology departments with an acute stroke unit. We were therefore able to test the transportability of the models, which constitutes accuracy in different

Received September 30, 2007; final revision received October 22, 2007; accepted October 24, 2007.

From the Institut für Medizinische Biometrie und Statistik (I.R.K., A.Z.), Universität zu Lübeck, Lübeck, Germany; Boehringer Ingelheim (E.B.), Ingelheim am Rhein, Germany; Neurologische Klinik (W.H.), Universität Heidelberg, Heidelberg, Germany; Stroke Trials Unit (P.M.W.B.), Division of Stroke Medicine, University of Nottingham, Nottingham, England; Miller School of Medicine (R.L.S.), University of Miami, Miami, Fla; and Klinik für Neurologie (H.C.D., C.W.), Universitätsklinikum Essen, Universität Duisburg-Essen, Essen, Germany.

Correspondence to Dr Inke R. König, Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany. E-mail inke.koenig@imbs.uni-luebeck.de

Stroke is available at http://stroke.ahajournals.org DOI: 10.1161/STROKEAHA.107.505867

Table 1. Estimated Regression Coefficients of the Original, Recalibrated, and Novel Models Predicting Functional Recovery and Survival

	Original	Recalibrated	Novel
Functional recovery			
Intercept	−5.782	−6.148	−5.112
Age	0.049	0.049	0.046
NIHSS*	0.272	0.272	0.196
Survival			
Intercept	−7.040	−7.373	−5.445
Age	0.049	0.049	0.037
NIHSS*	0.155	0.155	0.092

*Overall score on the NIHSS.

but similar populations.[18] The aim of the current study was to demonstrate more stringently the utility of our models by applying them to patients in the data set of the Virtual International Stroke Trials Archive (VISTA, www.vista.gla.ac.uk/).[19] Originating from diverse clinical trials in various countries, the stroke patients in VISTA differ from our original German Stroke Study data bank in terms of selection, level of stroke care, recruitment, and nationality. On the basis of our prognostic models, we will address the following questions: (1) Are the prognostic models adequate in this different population? (2) Can the predictions be relevantly improved by fine-tuning, ie, slight modifications, of the models? and (3) Can the predictions be relevantly improved by developing novel models, ie, deriving new parameter estimates?

Subjects and Methods

Model Development

A description of the development of the models has been detailed elsewhere.[17] In brief, functional independence of the patients was assessed by the Barthel Index (BI) as one of the most widely used measures of functional independence. To identify patients who recovered, as advocated in the guidelines for controlled clinical trials,[20] a cutoff BI value ≥95 versus <95 was used. Specifically, the following 2 models were developed with data from the stroke data bank of the German Stroke Collaborators[13,16,21]: (1) model I to predict functional recovery, ie, BI ≥95 versus BI <95 or dead and (2) model II to predict survival versus death (all causes).

From a set of 16 possible predictive variables that had been identified in a systematic literature search and included single items as well as the overall score of the NIHSS, logistic regression models were fitted by forward, backward, and stepwise selection. To model the relation with continuous variables, fractional polynomials were applied, and possible 2-way interactions were considered. The resulting logistic regression models showed a higher probability for the less favorable outcome in both models with higher age and greater overall neurologic impairment, as measured by the overall NIHSS score assessed within 6 hours after symptom onset (see Table 1, left column).[17]

To apply the resulting models, nomograms were created that provide the estimated probabilities for outcome and the resulting classifications based on age and NIHSS score of a single patient (see Figure 1).[22] To forecast the outcome for a specific patient, the values of each variable are marked on the respective lines. For instance, for the functional recovery model, age is marked on the second line of Figure 1A. Then, a straight line is drawn upward to determine the points for the variable "age." This is repeated for the NIHSS, and the points are summed and marked in the second to last line "Total Points." Drawing a line downward to the lowest line then gives the predicted probability for this patient to become functionally independent. For example, a patient aged 66 years (15 points) and with an overall NIHSS score of 7 (82 points) receives a total point score of 97. This corresponds to an estimated probability of ≈65% for functional independence. On the basis of the classification threshold from previous studies, this patient is therefore predicted to recover functionally. Similarly, application of the survival model (Figure 1B) results in 109 total points for this patient, corresponding to a 94% probability of survival.

VISTA Data Set

VISTA (www.vista.gla.ac.uk/) was created with the aim of providing access to patient data to perform exploratory analyses and hypothesis testing.[19] All included trials were approved by institutional review boards. At the time of data extraction (March 15, 2006), VISTA encompassed data from >15 000 patients from 21 acute stroke randomized, controlled trials. For the purpose of this analysis, relevant data were extracted from 11 trials that met the following entry criteria: (1) minimum data set of 100 patients; (2) documented entry criteria; (3) baseline assessment within 24 hours of stroke onset, including recording of neurologic deficit by NIHSS; (4) confirmation of stroke diagnosis by cerebral imaging within 7 days; (5) outcome assessed between 1 and 6 months after stroke onset, including recording of BI or mortality; and (7) monitoring procedures in practice to validate data. These data represented a total of 5843 patients who had been entered into VISTA after their inclusion in the respective clinical trials.

Statistical Analyses

For all analyses, patients with missing outcome data were excluded to allow for an evaluation of the resulting predictions. Patients in the original data set and in the VISTA data set are described with regard to age, NIHSS, and sex, and differences in the data sets are presented as mean differences and 95% CIs based on a *t* distribution of the difference (age and NIHSS) and as the difference in proportion with 95% CI, as proposed by Newcombe (method 10).[23]

To compare the applicability of our prognostic models, 3 different approaches were taken. In the most stringent approach, the algorithms as described in Figure 1 and Table 1 were applied to all patients in the VISTA data set for whom complete information on age and overall NIHSS score was available. The resulting numbers of correct classifications overall and in each outcome group were determined. In addition, a receiver operating characteristic was drawn for each model, which plotted specificity versus sensitivity, and the area under the curve is presented with its SE.

In the second approach, we aimed to adjust the original models to optimize the fit in the VISTA data set. For this, it should be remembered that a logistic regression model for prognosis principally consists of 2 components, the intercept and a set of slope coefficients. If the intercept is valid in the new data, the model is well calibrated. In contrast, if it is misspecified, the resulting predicted probabilities are systematically either too high or too low. On the other hand, if the slope is incorrectly estimated, the model shows insufficient discrimination in the new data, and the spread of the predicted probabilities is either too extreme or not extreme enough, so that the model cannot differentiate between patients with a more or less favorable outcome.[24] In this approach, we wanted to allow for a different prognosis but assumed that the effects of predictors would be similar. Therefore, we only aimed at recalibrating the models, ie, at adjusting the values of the intercept. This was solved technically by developing a logistic regression model to predict the observed outcomes in the VISTA data from the linear predictor of the original logistic regression model. The aims of this new regression model were to keep the slope fixed but to estimate the intercept. To meet this end, the linear predictor was used as an offset variable, thus fixing the coefficient at unity, so that the intercept was the only free parameter. The resulting estimate for the intercept indicates the deviation from the original one, and this model renders recalibrated predicted probabilities.[18] Only data from patients with complete information were used, and the resulting numbers of correct classifications overall and in each outcome group were determined.

Finally, novel logistic regression models were developed on the basis of the variables that had been selected for the previous models, namely, age and NIHSS score. The thresholds for categorization of patients were

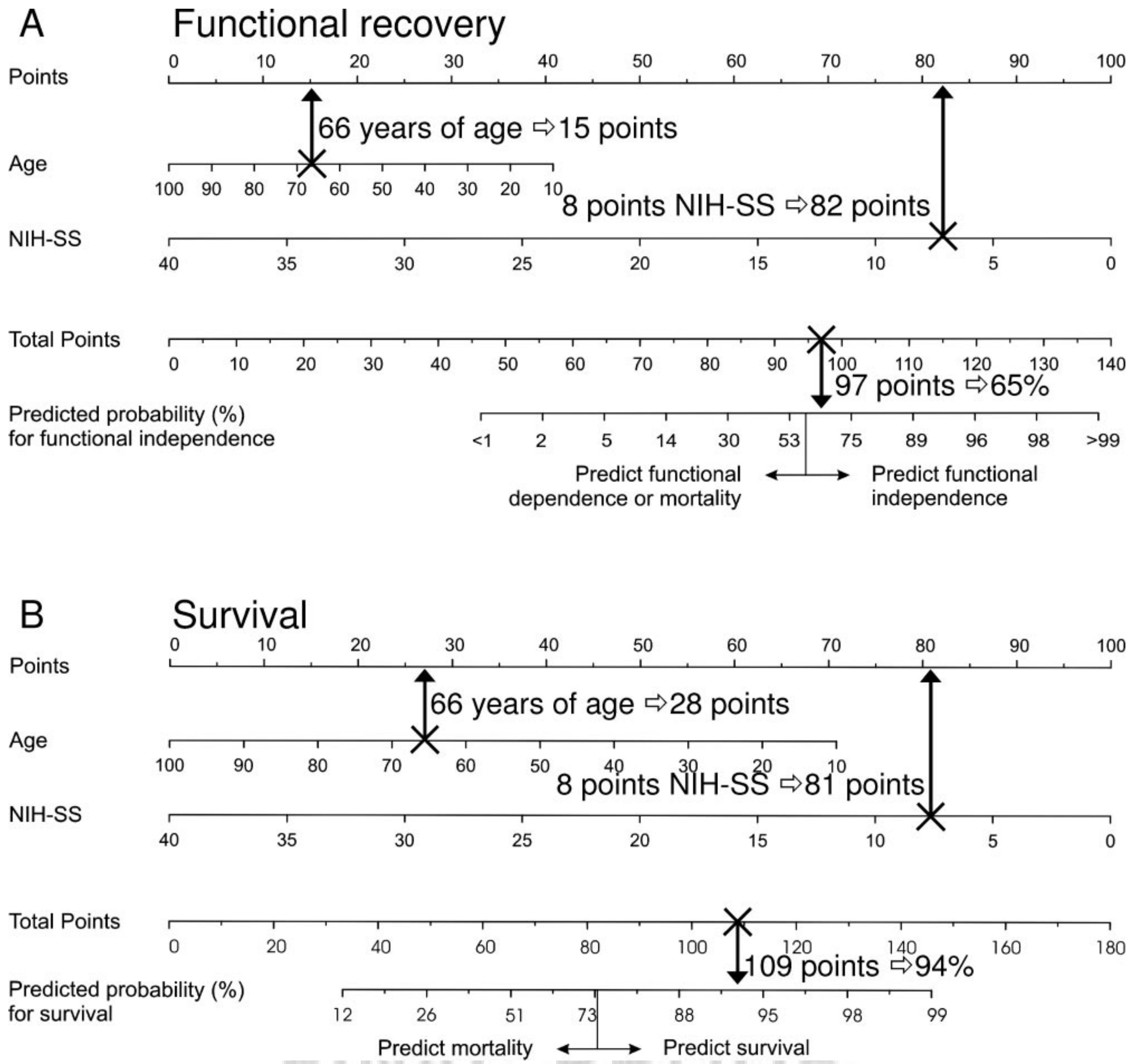

Figure 1. Nomograms for (A) the model predicting functional recovery (BI ≥95) vs no functional recovery (BI <95) or mortality and for (B) the model predicting survival vs mortality.

determined anew on the basis of the outcome frequencies in the data sets as before[17] to compare the resulting numbers of correct classifications with those from the previous approaches. Because in the different approaches the outcome of the same patients is predicted by different prognostic models, we tested differences in the overall accuracies by McNemar's test, and we present the estimated differences with 95% CIs according to Zhou and Qin.[25] The analyses were performed with the R software environment, version 2.3.1, with the Design package by Harrell.[26]

Results

Characteristics of the 5843 patients in the VISTA data set meeting the specified inclusion criteria are presented in Table 2. The BI after 90 days was recorded in 4441 patients, and 1970 of these had recovered, whereas 2471 had not functionally recovered. In another 978 patients without a recorded BI, information on mortality within this time frame was available. Of these, 607 (62.1%) had died, so that these were additionally classified as not functionally recovered. Therefore, for the functional recovery model, data from 5048 patients were available, of whom 1970 had recovered and 2471+607=3078 had either not recovered or had died. Independent of the availability of the BI, there was information on mortality for 5419 patients, of whom 4441 had survived and 978 were deceased. As shown in Table 2, age was slightly higher than in the original data sets (mean difference=0.71 years, 95% CI=0.04 to 1.38); similarly, neurologic impairment was less severe in the original sample (mean difference=6.50, 95% CI=6.15 to 6.84). In the VISTA data set, the proportion of women was higher than in the original data set (difference in proportions=3.4%, 95% CI=0.7% to 6.0%). Further details on patients' characteristics are given in the original publications.[17,19]

In the most stringent approach, the original models were applied to predict the patients' outcomes, and the receiver

Table 2. Patients' Characteristics in Original and VISTA Data Sets

	Original Data	VISTA Data
BI, n (%)		
≥95	1025 (58.4%)	1970 (44.4%)
<95	729 (41.6%)	2471 (55.6%)
Survival, n (%)		
Yes	1588 (90.5%)	4441 (82.0%)
No	166 (9.5%)	978 (18.0%)
Sex, n (%)		
Female	716 (40.8%)	2583 (44.2%)
Male	1038 (59.2%)	3260 (55.8%)
Age, y		
Mean (SD)	68.1 (12.7)	68.8 (12.3)
NIHSS*		
Mean (SD)	6.9 (6.2)	13.4 (6.5)

*Overall score on the NIHSS.

operating characteristics are shown in Figure 2. According to the original thresholds, 3678 patients (72.9%) were classified correctly overall in the functional recovery model, with a more correct prediction of patients who did not recover than in those who did (90.8% and 44.9%, respectively). According to the survival model, 3815 patients (70.4%) were classified correctly overall, with 56.3% of patients who died and 73.5% of patients who survived.

In the second approach, the original models were recalibrated by using an adapted intercept. The deviations of the original and the new intercepts were estimated to be 0.36 (95% CI=0.29 to 0.44) for the functional recovery model and 0.33 (95% CI=0.26 to 0.41) for the survival model, showing that the predicted probabilities for favorable outcome in both models were systematically too low (see Figures 3 A and 3B). Using the recalibrated models (see Table 1 for estimated regression coefficients and Figures 3C and 3D for calibration plots) led to slightly altered classifications, with overall 3735 patients (74.0%) being classified correctly for the functional recovery model (86.1% who did not recover and 55.1% who did recover). For the survival model, 4054 patients (74.8%) were predicted correctly (46.7% of patients who died and 81.0% of patients who patients). Thus, the accuracies were higher than in the original model (for the functional recovery model, difference=1.1%, 95% CI=0.4% to 1.9%, 2-sided $P=0.0026$, and for the survival model, difference=4.5%, 95% CI=3.7% to 4.5%, 2-sided $P=2.2\times10^{-16}$).

Finally, novel logistic regression models were developed by estimating new regression coefficients of the previously identified parameters (Table 1, right column). Thereby the recovery model predicted 3736 patients (74.0%) correctly overall (78.8% of those who did not recover and 66.6% of those who did). In the survival model, 4212 patients (77.7%) were classified correctly (38.2% of those who died and 86.4% of those who survived). Compared with the original model, the accuracy was higher overall (for functional recovery, difference=1.2%, 95% CI=0.1% to 2.3%, 2-sided $P=0.0452$; for survival, difference=7.4%, 95% CI=6.4% to 8.3%, 2-sided $P=2.2\times10^{-16}$). Although there was no difference in accuracies for the functional recovery model between

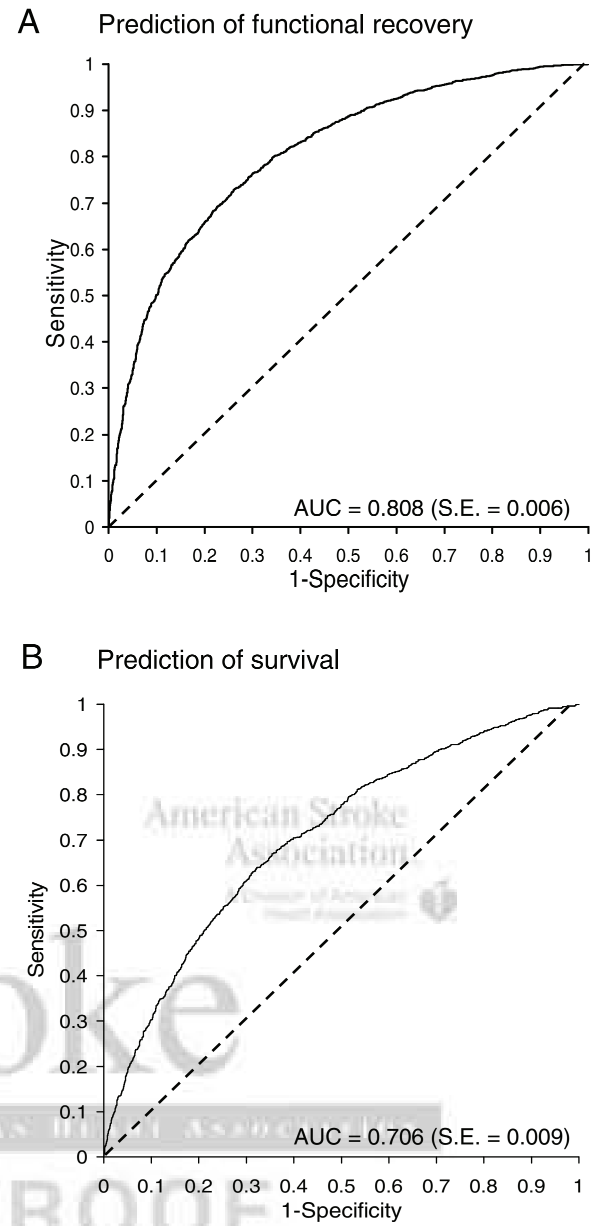

Figure 2. Receiver operating characteristics for (A) the model predicting functional recovery (BI ≥95) vs no functional recovery (BI <95) or mortality and for (B) the model predicting survival vs mortality based on the original model. Areas under the curves (AUC) are given with standard errors (S.E.).

the novel and the recalibrated prognosis (difference=0.0%, 95% CI=−0.8% to 0.9%, 2-sided $P=1.0000$), accuracy was higher in the novel than in the recalibrated survival model overall (difference=2.9%, 95% CI=2.3% to 3.6%, 2-sided $P=2.2\times10^{-16}$). However, this was mostly due to a more correct classification of the patients who survived, whereas <40% of patients who died were predicted correctly.

Discussion

An early, simple, and reliable model to calculate the prognosis of likely outcome in stroke patients is desirable and useful, for both clinical practice and research purposes. We previously developed and externally validated models that fulfilled

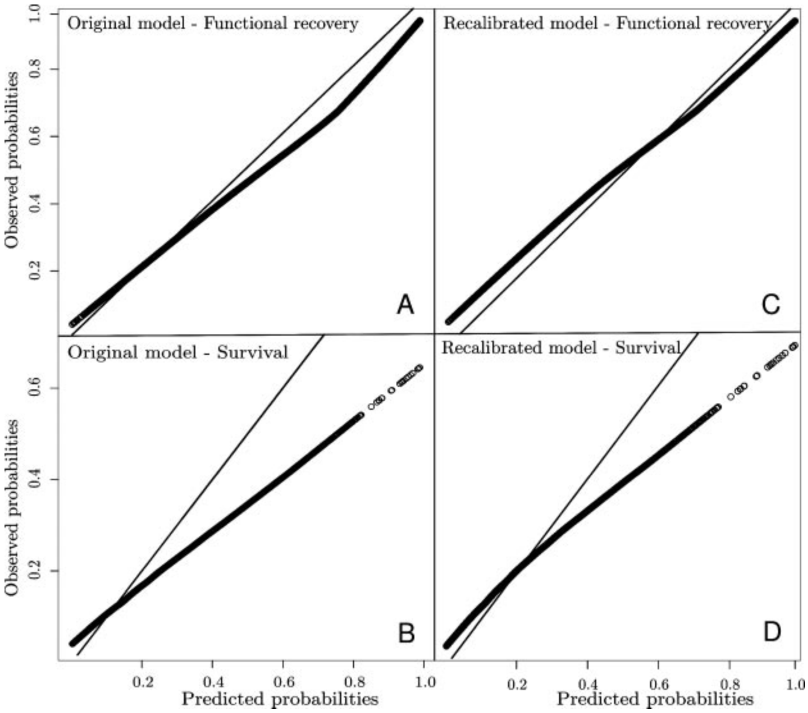

Figure 3. Calibration plots delineating observed vs predicted outcome probabilities from the original models predicting functional recovery (A) and survival (B) and from the recalibrated models predicting functional recovery (C) and survival (D). Dots represent the calibration curves created using lowess smoothers, and lines show the ideal calibration.

these criteria.[17] Starting from a set of 16 variables that had been identified in a systematic literature search, we devised a final model that included age and overall NIHSS score only, both as linear variables.[17] Thus, we generalized our already validated prognostic models in another phase III prognostic factor study by following the recommendations of Altman and colleagues.[27,28] This generalization aims at increased acceptance of the prognostic models by demonstrating their substantial predictive value in another population. Our prognostic models have been developed from the traditional statistical approach of logistic-regression models; these have the advantage of being simple to interpret and illustrate in a nomogram and have been shown to be at least as good in prediction as alternative machine learning approaches.[29]

In the present study, we have shown that our models, when transferred to more or less selected study populations, are able to correctly predict functional independence and mortality after 3 months in >70% of all patients. Because patients from controlled studies in VISTA were systematically predicted slightly too pessimistically by the original models, we performed a recalibration of the models in which the intercept of the model was adapted to the new data. As a result, the overall accuracy of prediction increased in the functionally independent or surviving patients but decreased for patients with a poorer outcome. This can be explained by the fact that the recalibrated models had a cutoff set according to the new outcome distribution rather than equal sensitivity and specificity. It needs to be emphasized that even though the refined models led to a significant increase of 57 and 241 patients being predicted correctly in comparison with the original models, the relative improvement in percentage is rather modest (≈1% and 4.5%, respectively). Also, with the adapted intercepts, the estimated accuracies of the refined models may be optimistic, so that in another independent sample, improvement over the original models may not stand.

Interestingly, estimating the coefficients anew does not improve the prediction over the recalibrated model for functional recovery, indicating a good discrimination of the original model. For the survival model, the higher overall accuracy of the new model was accompanied by a worse prediction of deceased patients, so that this is not recommended over the recalibrated model for clinical practice. To a greater extent than for the refined models, it should be considered that the latter models with novel parameter estimates need to be validated in different samples to guarantee generalization.

This study has some limitations. First, the predictive accuracies identified in the most stringent approach may not seem to justify relying on the given prediction over clinical judgment. However, we have previously shown that clinical judgment by the admitting neurology resident is inferior to our models and correctly predicted <70% of all patients.[17] Also, a simple recalibration of the intercept considerably improved the overall correct prediction of our models in VISTA. Second, our models do not consider imaging or laboratory investigations, which were impossible to obtain in our large original cohorts within an early time frame and a standardized evaluation protocol. Instead, we decided to focus our models on variables that are readily accessible and require neither a sophisticated technique nor a rigorous time frame. However, other studies have shown the prognostic value of magnetic resonance imaging in acute stroke,[9–11] which has also become an inclusion criterion in thrombolysis trials with desmoteplase.[30,31] Third, we have now shown the external validity of our models in 2 different populations of stroke patients, namely, patients admitted to German neurology departments with an acute stroke unit[17] and patients included in controlled clinical trials.[19] This by no means represents the entire universe of stroke patients, and subsequent studies are required to investigate the validity in other stroke populations.

Simple prognostic models may play an important role in future randomized trials in acute stroke. Patients included in these studies should have a high chance of incomplete recovery but low probability of mortality, which is usually unaffected by new medical treatment options. Therefore, it may be desirable to exclude patients with a high a chance of complete spontaneous recovery and those with a high chance of mortality because their data are unlikely to contribute to a measurable treatment effect. We have previously shown that overall sample size and trial time can be reduced by eliminating potential nonresponders and by increasing the number of eligible patients compared with conventional study designs.[3] Alternatively, prognosis-adjusted end points could be defined for patients with a high probability of functional recovery, as have already been used in several acute stroke trials.[32–34]

In conclusion, our original models can readily be applied in clinical practice and research settings with sufficient predictive accuracy, even in different patient populations. For patients included in clinical trials, a simple recalibration helps to adjust for a different case mix and is indeed recommended if a large data set is available.

Appendix

VISTA Steering Committee Members

VISTA steering committee members are as follows: K.R. Lees (chair), W. Hacke, R.L. Sacco, H.C. Diener, J. Grotta, P. Lyden, G.A. Donnan, S.M. Davis, P.M.W. Bath, N.G. Wahlgren, M. Hennerici, M. Kaste, M. Hommel, M. Fisher, S. Warach, J. Curram, P. Teal, B. Gregson, J. Marler, L. Claesson, and E. Bluhmki.

Disclosures

R.L.S. serves as a consultant and is on the Advisory Board for Boehringer Ingelheim. The remaining authors report no conflicts.

References

1. Counsell C, Dennis M, Lewis S. Prediction of outcome after stroke. *Lancet*. 2001;358:1553–1554.
2. Hand P, Wardlaw J, Lindley R, Keir S. Prediction of outcome after stroke. *Lancet*. 2001;358:1552–1553.
3. Weimar C, Ho T, Katsarawa Z, Diener H. Improving patient selection for clinical acute stroke trials. *Cerebrovasc Dis*. 2006;21:386–392.
4. Weir CJ, Kaste M, Lees KR. Targeting neuroprotection clinical trials to ischemic stroke patients with potential to benefit from therapy. *Stroke*. 2004;35:2111–2116.
5. Young FB, Lees KR, Weir CJ. Improving trial power through use of prognosis-adjusted end points. *Stroke*. 2005;36:597–601.
6. Wyatt JC, Altman DG. Commentary: prognostic models: clinically useful or quickly forgotten? *BMJ*. 1995;311:1539–1541.
7. Counsell C, Dennis M. Systematic review of prognostic models in patients with acute stroke. *Cerebrovasc Dis*. 2001;12:159–170.
8. Baird AE, Dambrosia J, Janket S, Eichbaum Q, Chaves C, Silver B, Barber PA, Parsons M, Darby D, Davis S, Caplan LR, Edelman RE, Warach S. A three-item scale for the early prediction of stroke recovery. *Lancet*. 2001;357:2095–2099.
9. Johnston KC, Connors AF Jr, Wagner DP, Haley EC Jr. Predicting outcome in ischemic stroke: external validation of predictive risk models. *Stroke*. 2003;34:200–202.
10. Johnston KC, Connors AF Jr, Wagner DP, Knaus WA, Wang X, Haley EC Jr. A predictive risk model for outcomes of ischemic stroke. *Stroke*. 2000;31:448–455.
11. Johnston KC, Wagner DP, Haley EC Jr, Connors AF Jr. Combined clinical and imaging information as an early stroke outcome measure. *Stroke*. 2002;33:466–472.
12. Counsell C, Dennis M, McDowall M, Warlow C. Predicting outcome after acute and subacute stroke: development and validation of new prognostic models. *Stroke*. 2002;33:1041–1047.
13. Weimar C, Ziegler A, König IR, Diener HC. Predicting functional and vital outcome after acute ischemic stroke. *J Neurol*. 2002;249:888–895.
14. Counsell C, Dennis M, McDowall M. Predicting functional outcome in acute stroke: comparison of a simple six variable model with other predictive systems and informal clinical prediction. *J Neurol Neurosurg Psychiatry*. 2004;75:401–405.
15. Counsell C, Dennis MS, Lewis S, Warlow C. Performance of a statistical model to predict stroke outcome in the context of a large, simple, randomized, controlled trial of feeding. *Stroke*. 2003;34:127–133.
16. The German Stroke Study Collaboration. Predicting outcome after acute ischemic stroke: an external validation of prognostic models. *Neurology*. 2004;62:581–585.
17. Weimar C, König IR, Kraywinkel K, Ziegler A, Diener HC. Age and the National Institutes of Health Stroke Scale within 6 h after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. *Stroke*. 2004;35:158–162.
18. König IR, Malley JD, Weimar C, Diener H-C, Ziegler A, on behalf of the German Stroke Study Collaboration. Practical experiences on the necessity of external validation. *Stat Med*. 2007;26:5499–5511.
19. Ali M, Bath PMW, Curram J, Davis SM, Diener HC, Donnan GA, Fisher M, Gregson BA, Grotta J, Hacke W, Hennerici MG, Hommel M, Kaste M, Marler JR, Sacco RL, Teal P, Wahlgren NG, Warach S, Weir CJ, Lees KR. The Virtual International Stroke Trials Archive (VISTA). *Stroke*. 2007;38:1905–1910.
20. Committee for Proprietary Medicinal Products (CPMP). Points to consider on clinical investigation of medicinal products for the treatment of acute stroke. European Agency for the Evaluation of Medicinal Products. 2000;CPMP/EWP/560/98.
21. König IR, Weimar C, Diener HC, Ziegler A. Vorhersage des Funktionsstatus 100 Tage nach einem ischämischen Schlaganfall: Design einer prospektiven Studie zur externen Validierung eines prognostischen Modells. *Z Arztl Fortbild Qualitatssich*. 2003;97:717–722.
22. Van Zee KJ, Manasseh DM, Bevilacqua JL, Boolbol SK, Fey JV, Tan LK, Borgen PI, Cody HS III, Kattan MW. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann Surg Oncol*. 2003;10:1140–1151.
23. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med*. 1998;17:873–890.
24. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol*. 2005;58:475–483.
25. Zhou X-H, Qin G. A new confidence interval for the difference between two binomial proportions of paired data. UW Biostatistics Working Paper Series. 2003; Report Number 205, University of Washington. Available at: http://www.bepress.com/uwbiostat/paper205.
26. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from http://www.R-project.org [21.12.2006].
27. Altman DG, Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer. *Breast Cancer Res Treat*. 1998;52:289–303.
28. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. *Br J Cancer*. 1994;69:979–985.
29. König IR, Malley JD, Pajevic S, Weimar C, Diener H-C, Ziegler A, on behalf of the German Stroke Study Collaborators. Patient-centered yes/no prognosis using learning machines. *Int J Data Mining Bioinformatics*. In press.
30. Furlan A, Eyding D, Albers G, Al-Rawi Y, Lees K, Rowley H, Sachara C, Soehngen M, Warach S, Hacke W. Dose Escalation of Desmoteplase for Acute Ischemic Stroke (DEDAS): evidence of safety and efficacy 3 to 9 hours after stroke onset. *Stroke*. 2006;37:1227–1231.
31. Hacke W, Albers G, Al-Rawi Y, Bogousslavsky J, Davalos A, Eliasziw M, Fischer M, Furlan A, Kaste M, Lees K, Soehngen M, Warach S. The Desmoteplase in Acute Ischemic Stroke Trial (DIAS): a phase II MRI-based 9-hour window acute stroke thrombolysis trial with intravenous desmoteplase. *Stroke*. 2005;36:66–73.
32. Krams M, Lees K, Hacke W, Grieve A, Porgogozo J-M, Ford G. Acute stroke therapy by inhibition of neutrophils (ASTIN): an adaptive dose-response study of UK-279,276 in acute ischemic stroke. *Stroke*. 2003;34:2543–2548.
33. Lees K, Zivin J, Ashwood T, Davalos A, Davis S, Diener H, Grotta J, Lyden P, Shuaib A, Hardemark H, Wasiewski W. NXY-059 for acute ischemic stroke. *N Engl J Med*. 2006;354:588–600.
34. Sherman D, Atkinson R, Chippendale T, Levin K, Ng K, Futrell N, Hsu C, Levy D. Intravenous ancrod for treatment of acute ischemic stroke: the STAT study: a randomized controlled trial. *JAMA*. 2000;283:2395–2403.