Decision tree analysis in subarachnoid hemorrhage: prediction of outcome parameters during the course of aneurysmal subarachnoid hemorrhage using decision tree analysis

Isabel Charlotte Hostettler, MD,1,2 Carl Muroi, MD,3 Johannes Konstantin Richter, MD,4,5 Josef Schmid, MA,6 Marian Christoph Neidert, MD,1 Martin Seule, MD,3,7 Oliver Boss, MD,3 Athina Pangalu, MD,4 Menno Robbert Germans, MD, PhD,1 and Emanuela Keller, MD1,3

Departments of 1Neurosurgery and 4Neuroradiology and 1Neurocritical Care Unit, Department of Neurosurgery, University Hospital Zurich; 3Department of Diagnostic, Interventional and Pediatric Radiology, University Hospital of Bern; 4Dynelytics, Zurich; 1Department of Neurosurgery, Kantonsspital St. Gallen, Switzerland; and 2Stroke Research Centre, University College London, Institute of Neurology, London, United Kingdom

OBJECTIVE The aim of this study was to create prediction models for outcome parameters by decision tree analysis based on clinical and laboratory data in patients with aneurysmal subarachnoid hemorrhage (aSAH).

METHODS The database consisted of clinical and laboratory parameters of 548 patients with aSAH who were admitted to the Neurocritical Care Unit, University Hospital Zurich. To examine the model performance, the cohort was randomly divided into a derivation cohort (60% [n = 329]; training data set) and a validation cohort (40% [n = 219]; test data set). The classification and regression tree prediction algorithm was applied to predict death, functional outcome, and ventriculoperitoneal (VP) shunt dependency. Chi-square automatic interaction detection was applied to predict delayed cerebral infarction on days 1, 3, and 7.

RESULTS The overall mortality was 18.4%. The accuracy of the decision tree models was good for survival on day 1 and favorable functional outcome at all time points, with a difference between the training and test data sets of < 5%. Prediction accuracy for survival on day 1 was 75.2%. The most important differentiating factor was the interleukin-6 (IL-6) level on day 1. Favorable functional outcome, defined as Glasgow Outcome Scale scores of 4 and 5, was observed in 68.6% of patients. Favorable functional outcome at all time points had a prediction accuracy of 71.1% in the training data set, with procalcitonin on day 1 being the most important differentiating factor at all time points. A total of 148 patients (27%) developed VP shunt dependency. The most important differentiating factor was hyperglycemia on admission.

CONCLUSIONS The multiple variable analysis capability of decision trees enables exploration of dependent variables in the context of multiple changing influences over the course of an illness. The decision tree currently generated increases awareness of the early systemic stress response, which is seemingly pertinent for prognostication.

https://thejns.org/doi/abs/10.3171/2017.7.JNS17677

KEY WORDS subarachnoid hemorrhage; decision tree analysis; clinical outcome; death; delayed cerebral infarction; shunt dependency; vascular disorders

C

LINICAL outcome in aneurysmal subarachnoid hemorrhage (aSAH) is significantly influenced by complications such as rebleeding, cerebral vasospasm, delayed cerebral ischemia (DCI), secondary infarction, and hydrocephalus. To predict complications and clinical outcome after aSAH, several clinical grading systems, e.g., Hunt and Hess;10 World Federation of Neurosurgical Societies (WFNS);25 HAIR (Hunt and Hess, age, intraventricular hemorrhage, rebleed);44 and FRESH (Functional Recovery Expected after Subarachnoid Hemorrhage);28 and radiological scores to predict cerebral vasospasm, such as Fisher grade5 and Barrow Neurological Institute (BNI) score,27 have been developed. However, their predictive accuracy remains limited. The complex pathogenesis and pleomorphic nature of the aforementioned complications certainly contribute to this fact. Nevertheless, the possibil-
ity of including prediction algorithms for specific complications in daily clinical assessments would allow crucial proactive decision making in the individual patient. The multivariable analysis capability of decision trees makes it possible to go beyond simple cause and effect relationships and to explore dependent variables in the context of multiple influences over time.\textsuperscript{10,22} The aim of this study was to create prediction models for outcome parameters using decision tree analysis based on easily accessible clinical and, in particular, laboratory data. In a decision tree analysis, patients are split in a recursive manner based on the values of a statistically important variable. The first splitting of the entire patient population into subgroups is based on the most important variable. The splitting is repeated, and the subsequent subgroups are further divided on its most important variable until the subgroup is no longer subject to further splitting.

**Methods**

**Database**

The database consisted of prospectively collected data from 548 patients with confirmed anSAH who were admitted to the Neurocritical Care Unit, University Hospital Zurich (an academic tertiary care center) from January 2007 to December 2013. In brief, the diagnosis was based on CT and catheter or CT angiography findings. In addition to patient's demographic characteristics, comorbidities were assessed (Table 1). The clinical severity at the time of admission was assessed according to the WFNS and Hunt and Hess grading systems.\textsuperscript{10,25} A neuroradiologist assessed the CT scans to determine the blood load (diameter of blood in a vertical cistern), Fisher grade, and BNI score, as well as the presence of infarcts. Delayed cerebral infarction was defined as radiologically or otherwise proven new infarcts, which did not occur within 48 hours of aneurysm coiling or clipping.\textsuperscript{4} We explicitly decided to define delayed cerebral infarction as an outcome variable as opposed to the more commonly used delayed cerebral ischemia (DCI), as the definition and use of this abbreviation has been inconsistent.\textsuperscript{26}

The database included laboratory data, from which the following parameters were selected for the decision tree analysis: glucose level at the time of admission, and interleukin-6 (IL-6), procalcitonin (PCT), C-reactive protein (CRP) levels, and leukocyte counts on days 1, 3, and 7 after hemorrhage. Glucose level on admission was selected because it has previously been described as a predictor for poor outcome after SAH.\textsuperscript{13,21} Hyperglycemia on admission was defined as a blood glucose level greater than 8 mmol/L. IL-6, PCT, CRP levels, and leukocyte counts were collected, as these systemic inflammatory parameters have been reported to correlate with occurrence of ischemic complications and/or poor clinical outcome.\textsuperscript{9,12,15,16,18} Patients underwent routine follow-up in the neurosurgical outpatient clinics 1 year after aSAH. The functional outcome was assessed according to the Glasgow Outcome Scale (GOS).\textsuperscript{11} The registry was approved by the ethics committee of Zurich, Switzerland.

**Statistical Analysis**

Continuous variables are presented as mean ± standard deviation and categorical variables as frequency and percentages. Statistical analysis was performed using IBM SPSS and IBM SPSS Modeler (versions 24 and 18, respectively).

**Selection of Predictors**

Univariable analysis was performed to screen candidate predictors regarding prediction of the following dependent variables: death, GOS score, occurrence of delayed cerebral infarction, and ventriculoperitoneal (VP) shunt dependency. All aforementioned clinical, radiological, and laboratory data were evaluated (Table 1). GOS scores were dichotomized as 1–3 (unfavorable outcome) and 4 and 5 (favorable outcome) for analysis. The importance of each variable was defined as \((1 - p)\), with \(p\) being the \(p\) value of the appropriate statistical test of association between the candidate predictor and the target variable. The Pearson’s chi-square test was used for categorical predictors, and the continuous 1-way ANOVA F-test was used for continuous variables. To identify the most important predictors to be included in the model, the respective \(p\) values were subsequently ranked in ascending order.

**Decision Tree Development and Internal Validation**

To examine the model performance, the cohort was randomly divided into a derivation cohort (60% [\(n = 329\)], training data set) and a validation cohort (40% [\(n = 219\]), test data set). The classification and regression tree prediction algorithm was applied to predict death, dichotomized functional outcome, and VP shunt dependency. Chi-square automatic interaction detection was applied to predict delayed cerebral infarction.\textsuperscript{2} The prediction model for each of the outcome parameters was adapted for the specific time points of days 1, 3, and 7 after aSAH.

**Missing Values**

If the dependent variable of a case was missing, the case was ignored in the analysis. If all predictor variables of a case were missing, the case was ignored. With the classification and regression tree algorithm, the surrogate split method was otherwise used to deal with missing data in predictor variables. With the chi-square automatic interaction detection algorithm, missing values were treated as a predictor category. Based on the smallest \(p\) value, the algorithm decided whether to merge the missing category with its most similar category or to keep the missing category as a separate category.\textsuperscript{5}

**Model Presentation**

For each of the dependent variables, a risk chart was created with regard to the absence or presence of the independent variables. The group containing all patients is termed “root,” and the subgroups are termed “nodes.” Starting at the root, an appropriate primary split was selected, after which the data set was further divided into smaller subsets at each node up to the point where no further information could be gained or the stopping criterion was reached. The most significant independent variable tested is indicated below the root or node. In the branches to the next node, corresponding threshold values of this independent variable can be found. The improvement of the model with each additional layer can be estimated by the numerical value below the label of the independent variable. The performance of each model was assessed by comparing its
sensitivity and specificity in predicting a specific event in both the training and test data sets. The quality of prediction was estimated to be good if the difference between the training and test data sets was < 5% and acceptable if it was < 10%. Based on the recorded factors of a new patient, risk can be predicted by following the decision tree path, beginning from the root to the nodes.

Results
The mean age of all patients was 54.7 ± 13.25 years (range 14–87 years); 368 patients (67.2%) were female and 180 (32.8%) male. Of the population, 226 (41.2%) patients sustained a high-grade SAH (WFNS grades III–V), and 320 (58.4%) a low-grade SAH (WFNS grades I and II). Two hundred eighty-eight (52.5%) patients were treated with clipping, 214 (39.0%) with coiling, 13 patients had a combination of coiling and clipping (2.4%), 2 patients received a bypass for trapping of the aneurysm (0.4%), and 31 patients (5.7%) were treated conservatively. Patients’ baseline characteristics are shown in Table 1.

Prediction of Outcome
Prediction of Death
A total of 101 of 548 patients died, resulting in an overall mortality of 18.4% (Table 2). The decision tree models predicting death or survival on days 1, 3, and 7 after aSAH are shown in Figs. 1–3. The most important differentiating factor on day 1 was the IL-6 level. The pathway with

CONTINUED FROM PREVIOUS COLUMN

TABLE 2. Outcome characteristics (dependent variables)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. of Patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall mortality</td>
<td>101 (18.4)</td>
</tr>
<tr>
<td>Outcome after 1 yr</td>
<td></td>
</tr>
<tr>
<td>GOS scores 1–3 (unfavorable)</td>
<td>137 (31.4)</td>
</tr>
<tr>
<td>GOS scores 4 &amp; 5 (favorable)</td>
<td>300 (68.6)</td>
</tr>
<tr>
<td>Delayed infarction</td>
<td>127 (23.2)</td>
</tr>
<tr>
<td>VP shunt dependency</td>
<td>148 (27.0)</td>
</tr>
</tbody>
</table>
FIG. 1. Decision tree models on day 1 for the occurrence of death. Lc = leukocytes.
FIG. 2. Decision tree models on day 3 for the occurrence of death.
the best prediction of death was as follows: IL-6 on day 1 greater than 10.35 ng/L (node 2), leukocyte counts on day 1 greater than $11.7 \times 10^9$/L (node 6), and glucose on admission greater than 9.4 mmol/L (node 14). The model had the greatest accuracy on day 1. The sensitivity to predict death was 83.1% in the training and 60% in the test data set, and the specificity was 75.3% and 71.0%, respectively (Table 3). The most important differentiating factor on days 3 and 7 was the leukocyte count on day 1. The specificity was 83.0% in the training and 76.7% in the test data set on day 3, and 78.2% in the training and 71.6% in the test set on day 7.

FIG. 3. Decision tree models on day 7 for the occurrence of death.
I. C. Hostettler et al.

Prediction of Functional Outcome

The outcome data regarding GOS score after 1 year were available in 437 patients. Favorable functional outcome, defined as GOS scores 4 and 5, was observed in 68.6% of patients (Table 2). For the dichotomized outcome (GOS scores 4 and 5 as favorable outcome; GOS score ≤3 as unfavorable outcome), the decision tree demonstrated the same prognostic subgroups and ranking order at all time points (Fig. 4). The pathway with the best prediction of unfavorable outcome was PCT on day 1 of greater than 0.23 μg/L and of favorable outcome PCT on day 1 of 0.23 μg/L or lower, as the most important differentiating factor (node 1), followed by WFNS grade II or lower (node 3). At each time point, the prediction accuracy for favorable outcome was 71.1% in the training and 66.7% in the test data set (Table 3), and the prediction accuracy for unfavorable outcome was 82.9% in the training and 75.4% in the testing data set.

Prediction of Delayed Cerebral Infarction

A total of 127 patients (23.2%) developed delayed cerebral infarction. The decision tree revealed the same prognostic subgroups and ranking order at all time points (Fig. 5). A CRP level on day 1 of 23 mg/L or lower was the most important differentiating factor. The pathway with the best prediction of absent delayed cerebral infarction was CRP level on day 1 of less than 23 mg/L (node 1), CRP level on day 1 of 11 mg/L or lower (node 3), no present comorbidi-

![FIG. 4. Decision tree models for the dichotomized outcome were the same on days 1, 3, and 7. GOS scores of 1–3 were defined as “unfavorable” and GOS scores of 4–5 as “favorable.”](image)
ties (node 7), and PCT level on day 1 of 0.26 µg/L or lower (node 13). The prediction accuracy for the absence of cerebral infarction was 79.7% in the training and 69.4% in the test data set at each time point (Table 3).

Prediction of Shunt Dependency
A total of 148 patients (27%) developed VP shunt dependency. The same decision tree fit best for the first 2 time points (Fig. 6). The best prediction was achieved with the decision tree on day 7 (Fig. 7). The most important differentiating factor was hyperglycemia on admission (root), followed by the localization of the aneurysm (node 1) and WFNS grade (node 2). Overall, 89.8% of the patients with absence of hyperglycemia on admission, ruptured aneurysm in the anterior circulation, and no concomitant malignant disease did not develop VP shunt dependency. Regarding the absence of VP shunt dependency, the model was most accurate on day 7 with a prediction accuracy of 87.3% in the training set and 80.9% in the test data set (Table 3).
Discussion

The accuracy of the decision tree models was good for survival on day 1 and for favorable functional outcome at all 3 time points, with a difference between the training and test data set of less than 5%. Prediction accuracy for survival on day 1 was 75.3% in the training data set. The most important differentiating factor was the IL-6 level on day 1. Systemic IL-6 levels have previously been reported to correlate with the clinical severity grade, occurrence of delayed neurological deficits, and outcome after aSAH.\textsuperscript{16,17} However, in those studies, serial measurements over the course illness were averaged and analyzed. Less is known about the predictive value of early (≤ 24 hours) levels. The IL-6 level on day 1 might reflect the severity of the initial stress response in a selective manner.

Favorable functional outcome at all time points had a prediction accuracy of 71.1% in the training data set, with PCT on day 1 being the most important differentiating factor at all time points. PCT, as an acute-phase protein, is elevated under various other acute stress conditions, such as polytrauma, cardiac arrest, and burns.\textsuperscript{1} It can be assumed that PCT levels in aSAH patients reflect an acute systemic stress response to the hemorrhage. Of further interest, early (≤ 24 hours) PCT levels have been reported to accurately predict unfavorable neurological outcome after transient global cerebral ischemia due to cardiac arrest.\textsuperscript{6,8} In severe aSAH, it is known that comparable transient global cerebral ischemia occurs.\textsuperscript{3}
For survival, the prediction accuracy was acceptable on days 3 and 7 (difference between training and test data set < 10%). The prediction accuracy for survival on day 3 was 83% and 78.2% on day 7. In both cases, a high leukocyte count on day 1 was the most important differentiating factor. Leukocytosis during the course of the illness has been linked with occurrence of vasospasm, DCI, and/or outcome. More recently, initial (≤ 24 hours) leukocyte counts have been reported to be predictive for DCI and poor neurological outcome. The most important differentiating factor for shunt dependency was hyperglycemia on admission. This corresponds to the evaluation of the Columbia University database (published in 2010). Glucose values of at least 126 mg/dl (7 mmol/L) were found to be associated with long-term shunt dependency.

The present study first analyzed predictors for outcome parameters in a time-dependent manner on days 1, 3, and 7. Interestingly, laboratory parameters obtained on day 1 were the most important differentiating factors in our statistical model and thereby surpassed conventional scoring systems, such as the Hunt and Hess and WFNS grading systems. This might be explained by the fact that these grades are assigned by humans and thus are subject to interrater variability, whereas laboratory parameters are absolute values. The results underline the significance of the very initial systemic stress response, reflecting the impact of the hemorrhage on the entire organism, for the overall mortality and morbidity.

In clinical practice, decision making is required on a daily and continuous basis. Even more, critical threshold values of laboratory parameters are not static but change frequently during the course of a disease, as can be seen.

FIG. 7. Decision tree models for VP shunt dependency on day 7. ACA = anterior cerebral artery; AcomA = anterior communicating artery; AICA = anterior inferior cerebellar artery; BA = basilar artery; ICA = internal carotid artery; MCA = middle cerebral artery; PcomA = posterior communicating artery; PericalA = pericallosal artery; VA = vertebral artery.
in our data set regarding survival prediction on days 3 and 7. In our survival prediction model, leukocyte count on days 1, 3, and 7 was consistently the most important differentiating factor, but the threshold varied from 12.33 × 10^9/L or lower on day 3 to 13.02 × 10^9/L or lower on day 7. Especially after aSAH, secondary inflammatory as well as reparatory processes change the requirements of the organism over the course of time. Threshold values according to which a patient is treated have to be adapted on a daily basis. Prediction scores have to take into account continuous changes in a patient’s requirements depending on the patient’s characteristics and current laboratory and clinical parameters, as well as complications that arise during the course of the disease, while providing an accurate prediction of the outcome variable of interest.

Our study has some limitations. Decision tree analysis is of limited use when missing data are present, which commonly occurs in medical studies. As our data set contained variables with missing values, only algorithms with well-defined criteria regarding the handling of missing values could be used. Another drawback regarding a decision tree is its difficult application on multifactorial, complicated diseases such as aSAH, as the more nodes there are, the less accurate the expected outcome is. Also, decision trees are based on expectations, which can lead to errors. A further clear limitation of the current study is the relatively low number of patients, and the fact that an external validation has not been performed. This is reflected in the restricted sensitivity for the specific event and the large differences between the training and test data sets as in the prediction of delayed cerebral infarctions. Therefore, it is of great interest to target these limitations with a harmonized large multicenter database to 1) perform an external validation and 2) increase the power and sensitivity of the analysis.

Despite the stated limitations, we have created a starting point for future large, multicenter studies to develop more sensitive and specific prognostic scores in patients with aSAH.

Conclusions

Prediction scores are powerful tools regarding the management and decision-making process in general and especially in patients with complex, multifactorial diseases, such as aSAH. The current study demonstrates the potential benefit of decision tree analysis in patients with aSAH. The necessity to handle an increasing amount of information calls for tools that can process and analyze large amounts of data, i.e., from continuously measured parameters over the course of an illness. Decision tree analysis is an interesting tool that can be used in data mining to generate new information based on an existing database, such as the one presented. The decision tree currently generated increases awareness of the early systemic stress response, which is seemingly pertinent for prognostication. To increase its power, sensitivity, and specificity, its accuracy needs to be validated in future studies of large cohorts.

References

Disclosures
The Helmut Horten and the Gebert Ruef Foundations funded this study.

Author Contributions
Conception and design: Hostettler, Schmid, Germans, Keller. Acquisition of data: Hostettler, Muroi, Richter, Neidert, Boss, Pangalu, Keller. Analysis and interpretation of data: Hostettler, Neidert, Keller. Drafting the article: Hostettler, Richter, Neidert, Keller. Critically revising the article: Hostettler, Muroi, Richter, Neidert, Seule, Pangalu, Germans, Keller. Reviewed submitted version of manuscript: Hostettler, Muroi, Schmid, Neidert, Seule, Boss, Pangalu, Germans, Keller. Approved the final version of the manuscript on behalf of all authors: Hostettler. Statistical analysis: Hostettler, Schmid. Administrative/technical/material support: Keller. Study supervision: Keller.

Supplemental Information
Previous Presentations
Portions of this work were presented as a talk at the Annual Conference on Intracranial Pressure, Brain Perfusion, and Hydrocephalus, German Society for Neurosurgery (DGNC), Bern, Switzerland, November 2016; and at the Annual Conference on Intensive Care and Neuro-Traumatology, German Society for Neurosurgery (DGNC), Mannheim, Germany, February 2017.

Correspondence
Isabel Hostettler: University College London, United Kingdom. isabel.hostettler@gmail.com.