Published in final edited form as:

World Neurosurg. 2017 May ; 101: 577–583. doi:10.1016/j.wneu.2017.02.082.

Prognosticating Functional Outcome Following Intracerebral Hemorrhage: The ICHOP Score

Vivek P. Gupta, B.A.¹, Andrew L.A. Garton, B.S.¹, Jonathan A. Sisti, B.S.¹, Brandon R. Christophe, B.A.², Aaron S. Lord, M.D.³, Ariane K. Lewis, M.D.³, Hans-Peter Frey, Ph.D.⁴, Jan Claassen, M.D., Ph.D.⁴, and E. Sander Connolly Jr., M.D.²

¹College of Physicians and Surgeons, Columbia University (30 W. 168th St., Suite 5-454 New York, NY, 10032)

²Department of Neurosurgery, Columbia University, College of Physicians and Surgeons (30 W. 168th St., Suite 5–454 New York, NY, 10032)

³Departments of Neurology and Neurosurgery, Division of Neurocritical Care NYU Langone Medical Center (530 First Ave, New York, NY, 10016)

⁴Division of Critical Care Neurology, Department of Neurology, Columbia University, College of Physicians and Surgeons (177 Fort Washington Avenue, MHB 8 Center, Room 300, New York, NY, 10032)

Abstract

Background—The morbidity, mortality, and monetary cost associated with intracerebral hemorrhage (ICH) is devastatingly high. Several scoring systems have been proposed to prognosticate outcomes following ICH, though the original ICH Score is still the most widely used. However, recent research suggests that systemic physiological factors, such as those included in the APACHE II score, may also influence outcome. Additionally, no scoring systems to date include pre-morbid functional status. Therefore, we propose a scoring system that incorporates these factors to prognosticate 3- and 12-month functional outcomes.

Methods—We used the Random Forest machine learning technique to identify factors from a dataset of over 200 data points per patient that were most strongly affiliated with functional outcome. We then used linear regression to create an initial model based on these factors and modified weightings to improve accuracy. Our scoring system was compared to the ICH Score for prognosticating functional outcomes.

Results—Two separate scoring systems (ICHOP₃ and ICHOP₁₂) were developed for 3- and 12-month functional outcomes using GCS, NIHSS, APACHE II, pre-morbid modified Rankin scale (mRS), and hematoma volume (3-month only). Patient outcomes were dichotomized into good

Correspondence for this paper should be sent to Vivek Gupta at 30 W. 168th St., Suite 5-454 New York, NY, 10032, Telephone: (212)-305-4679 sean.vivek.gupta@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The authors declare no conflicts of interest.

(mRS 0–3) and poor (mRS 4–6) categories based on functional status. AUCs in the derivation cohort for predicting mRS were 0.89 (3-month) and 0.87 (12-month); both were significantly more discriminatory than the original ICH Score.

Conclusion—The ICHOP scores may provide more comprehensive evaluation of a patient’s long-term functional prognosis by taking into account systemic physiological factors as well as pre-morbid functional status.

Keywords

functional outcome; ICH Score; intracerebral hemorrhage; prognostication; stroke

Introduction

Intracerebral hemorrhage (ICH) accounts for 10–15% of all strokes in the United States.¹ The loss of life associated with ICH is high, with approximately 35% mortality at 30 days and mortality at 1 year reaching as high as 59%.² ICH morbidity is also staggering: one study put the aggregate lifetime cost of ICH at \$6 billion.³ United States hospital admissions for ICH have increased by over 15% during the 2000s, possibly due to the increasing age of the population.⁴

Over the past decade, many scoring systems have been created with the intention of prognosticating outcomes for ICH patients, beginning with the original ICH Score in 2001 and followed by other scales such as the Essen ICH Score, ICH-FOS, FUNC score, and AVICH score.^{5,6,7,8,9} Although many of these demonstrate acceptable performance in prognosticating outcome, each has its own associated problems: limited external validity, not being designed to predict functional outcome, or limited testing in American populations.^{10,11,12,13} Due to its simplicity and validity throughout multiple cohorts, the ICH Score remains the most commonly used grading scale for ICH mortality and functional outcome.^{6,14}

Most of these grading systems have similar makeup, including factors such as Glasgow Coma Scale (GCS) or National Institutes of Health Stroke Scale (NIHSS), hematoma volume and location, and age. Recently, however, it has been suggested that other physiological factors may play an important role in predicting prognosis following ICH, particularly those included in the APACHE II score.¹⁵ Additionally, most scales use either the GCS or the NIHSS to assess neurological function, though each is subject to its own problems; for example, the NIHSS score may be identical in patients with different stroke volumes in different hemispheres.^{16,17} Finally, to our knowledge, no published scores to date account for the previous functional status of the patient prior to ICH, which has not yet been explored as a predictor of long term functional outcome. It therefore follows that inclusion of robust physiological admission data, multiple methods of neurological assessment, and patient baseline may produce a more accurate and comprehensive prediction algorithm for functional outcome when combined with measures already utilized by previous scoring systems.

Methods

Study Population and Data Collection

Data in the Columbia University Intracerebral Hemorrhage Outcomes Project (ICHOP) database (Columbia University IRB-AAAD4775) was collected prospectively by trained clinicians and researchers from patients presenting to Columbia University Medical Center (CUMC) with a diagnosis of intracerebral hemorrhage from March 2009 through May 2016 (n=575). Data collected includes patient identification information, admission scores, admission labs, clotting information, etiology, procedures performed, treatments provided, neurological deterioration during stay, hospital complications, CT scan information, past medical history, and discharge data. Trained laboratory personnel collected follow-up information through phone interviews with patients or caregivers. Patients were excluded from the final derivation cohort if they lacked one or more of the necessary components of admission or follow-up information included in the model. Out of the original 575 patients in the database, there were 365 and 321 patients with complete data for our 3-month score and 12-month score, respectively. Thus, it is from these cohorts that the two scores were derived.

The study protocol was approved by the institutional review board (IRB) at CUMC. Written informed consent was obtained from patients or family members if patients were unable to consent themselves.

Outcome Reporting

We used the modified Rankin scale (mRS) as our sole outcome measure.¹⁸ The mRS is used to assess functional disability in neurologically impaired patients, primarily following a stroke; it is measured from 0 (no disability) to 6 (death), with each subsequent level indicating progressive functional impairment.¹⁸ We examined mRS scores at 3 months and 12 months post-discharge in our study. Finally, we dichotomized mRS scores into good functional outcome (mRS 0–3) and poor functional outcome (mRS 4–6), as this method allowed for a stronger classification system, easier comparisons to other systems, and simpler interpretation of results; this re-classification approach has demonstrated efficacy before.¹⁹

Model Formulation and Refinement

Random Forest is a machine learning technique that uses partial decision tree-based analyses to classify subjects into categories based on training datasets with similar characteristics. In doing so, it evaluates the variables in its training datasets that are most useful in classifying new patients by reporting how much decision-making error is reduced by the inclusion of each factor. We were therefore able to use this technique (Python 2.7; SciPy stack 0.18.0; `sklearn.ensemble.RandomForestClassifier`) to determine which pieces of information had the greatest utility in classifying, or prognosticating, outcomes. We refer to the ability of a variable in classifying outcomes as its “importance” for this reason. Additionally, when Random Forest is used with a binary classifier label (the prognosis, or outcome), it is able to report the area under the receiver operating characteristic curve (AUC), a proxy for the strength of a classification system that utilizes false negative and false positive readings. In a

stepwise fashion, we were able to compare AUCs between models to eliminate variables that did not significantly increase the strength of our model.

We used any information with a normalized “importance” rating of 0.01 or greater (i.e. normalized reduction in error of 1% or more) as the foundation of our model, excluding any information that was not available upon admission. We then used least-squares linear regression (Python 2.7; SciPy stack 0.18.0; sklearn.linear_model.LinearRegression) to create an initial model for our scoring systems. We found that coefficients derived from total explained variance, rather than beta values, created a more accurate initial model in both our 3-month and 12-month scores, termed ICHOP₃ and ICHOP₁₂, respectively. In order to streamline and simplify the model, factors that did not contribute significantly to the model were removed in a step-wise fashion.

Of note, to eliminate redundancy in clinical variables, APACHE II was converted into APACHE II Physiologic score which does not include the GCS component, as GCS was already included in our model.²⁰ Additionally, age was removed, as it was accounted for by the APACHE II score and did not significantly contribute on its own. For ease of calculation and simplicity, GCS, NIHSS, and hematoma volume were converted from continuous variables into categorical data based on groupings found in previous studies, as shown in Table 2.^{7,21} These categories were then assigned a value for use in the model. Finally, ICH Score was not included in the model, as it was the scoring system the model was being compared against.

External Validation

We calculated ICHOP₃, ICHOP₁₂, and ICH Scores for an external cohort of 97 patients at an unaffiliated center and compared their ability to predict good or poor functional outcome as shown by their AUCs. ICHOP scores were calculated externally with a pre-formatted spreadsheet, and de-identified ICHOP and ICH Scores versus functional outcome were then compared.

Statistical Analysis

We used Python (version 2.7), R (version 3.3.0) and Microsoft Excel (version 15.24) to perform statistical analysis and organize presentation of data. We then calculated AUC for our Random Forest model, linear regression model, and final adjusted model to assess sensitivity and specificity for prediction of poor outcome, though only the AUC information for the final adjusted models was used for evaluation and comparison purposes. To compare AUCs, we used the R package pROC (version 1.8) to implement the method described by DeLong and colleagues as well as calculate AUC confidence intervals.²² We used the same software to conduct post-hoc power analyses. We also calculated Cox/Snell, McFadden, and Nagelkerke R² values to evaluate goodness of fit for determining binary functional outcome.

Results

Of 575 patients who presented to CUMC with ICH between March 2009 and May 2016, 365 had complete information at 3 months and, of these, 321 also had complete information at 12 months. Table 1 shows cohort data for patients with complete information at 3 months.

This derivation cohort was similar to the database in terms of important factors such as age, GCS, NIHSS, and hematoma volume.

From our available data, the factors available upon admission with the highest Random Forest “importances” for predicting mRS at 3 months included age, GCS, NIHSS, APACHE II, ICH Score, several ICH Score variations (e.g. Essen ICH Score), hematoma volume, and premorbid mRS scores. Other factors examined via Random Forest were either not available at admission, available for too few subjects, or showed significantly lower “importance” ratings.

The initial linear regression gave an R^2 value of 0.45. The amount of variance explained by each factor formed the basis for the weights assigned to each variable. Although most factors remained relatively constant from their initial derivation to the final model, Random Forest “importances” and beta coefficient value ratios suggested that GCS weighting should be higher than the value originally ascribed, ultimately improving the accuracy of the model. Our final model to predict functional outcome at 3 months is shown below.

$$ICHOP_3 \text{ Score} = \frac{20 * \{GCS\} + 3 * \{NIHSS\} + APACHE II_{phys} + 1.5 * \{Hematoma Volume\} + 8 * \text{premorbid mRS}}{11}$$

The score is rounded to the nearest whole number for comparison purposes. Bracketed factors are assigned a new value based on their categorical grouping before being entered into the score (see Table 2). APACHE II Physiologic score and premorbid mRS scores are entered without conversion. The score was divided by 11 to provide an eight-point final scale.

We then applied this approach to create a model that predicted functional outcome after 12 months using the same cohort, excluding those that did not have 12-month mRS scores for a total of 321 patients. Using the same five factors, our initial linear regression had an R^2 value of 0.44. As before, we used the amount of variance explained by each factor as the basis for our weightings. However, we found that hematoma volume explained an insignificant amount of variance compared to the other four factors and therefore dropped it from this version of the model, which is shown below.

$$ICHOP_{12} \text{ Score} = 3 * \{GCS\} + \{NIHSS\} + 2 * APACHE II_{phys} + 4 * \text{premorbid mRS}$$

Additionally, the ICHOP12 score can be calculated via a table (Table 3).

When applied to the derivation cohort, the ICHOP₃ and ICHOP₁₂ scores stratify patients by functional outcome at 3 months and 12 months (see Figure 1). In ICHOP₃, patients with scores 7 and above were categorized into a single group, as the proportion of patients with poor outcomes were nearly identical in each. ICHOP₁₂ patients were stratified into categories by blocks of 15. Similarly, patients with scores above 90 were categorized into a single group because 100% of them exhibited a poor outcome at 12 months.

In our derivation cohort, the ICHOP₃ score stratified patients into 8 groups where poor 3-month outcomes were observed in 0, 12, 27, 51, 66, 80, 92, and 96% of patients. The ICHOP₁₂ score stratified patients into 7 groups with poor 12-month outcomes in 0, 37, 60, 74, 83, 94, and 100% of patients (Figure 1). Performance measures for discrimination and goodness-of-fit for ICHOP₃, ICHOP₁₂ scores, and corresponding ICH Scores (for comparison purposes) are shown in Table 4. AUCs of ICHOP₃ and ICHOP₁₂ scores in the derivation cohort are shown in Figure 2. DeLong's method for comparison of AUCs showed that ICHOP₃ and ICHOP₁₂ scores were significantly better than ICH Scores for discriminating between favorable and poor functional outcomes (3-month $p = 0.046$; 12-month $p = 0.032$).²²

In our external validation cohort of 97 patients, we found that the ICHOP₃ score was more discriminatory than the ICH Score for predicting functional outcome at 3 months (AUC 0.745 vs. 0.691); however, this difference was not statistically significant ($p = 0.63$). In the same cohort, we also found that the ICHOP₁₂ score was more discriminatory than the ICH Score for predicting outcome at 12 months (AUC 0.843 vs. 0.664), though the difference was still not statistically significant ($p = 0.15$).

Discussion

Here we present a novel scoring system for predicting functional outcome following ICH at 3 and 12 months follow-up. Both models show robust discrimination as determined by AUC (ICHOP₃: 0.89; ICHOP₁₂: 0.87) and are demonstrated to be significantly better than the original ICH Score in predicting 3- and 12-month functional outcome in our derivation cohort. Several ICH scoring systems have been developed since the induction of the original ICH Score in 2001.^{5,6,7} The ICHOP scores take into account factors used by other ICH scoring systems (GCS, NIHSS, and hematoma volume for 3-month outcome) while also incorporating novel factors such as pre-morbid functionality and APACHE II score to develop a more extensive model for prognosticating functional outcome. The goal of this study was not to create a replacement score for the widely-used ICH Score or any of its derivatives but instead to supplement them with a more precise method of determining long-term functional outcome.

Incorporation of the APACHE II score is particularly relevant given new research that demonstrates that this score may have prognostic validity for ICH patients.¹⁵ Although that particular study focused on mortality, machine-learning analysis in our study demonstrated that APACHE II score also was an important predictor of functional mortality. Given the acutely ill status of ICH patients, including a scoring system that takes into account a wide variety of physiological factors may help to provide a more comprehensive picture of an ICH patient. Unfortunately, the complexity of calculating the APACHE II score may be a barrier to its use when compared to more easily accessible scales such as the original ICH Score, particularly given that both are used to assess mortality at 30 days. However, in critical care environments where APACHE II is routinely calculated, incorporating it into the ICHOP scoring system is, in principle, a relatively simple task.

Because the original ICH Score is the most commonly used and easily accessible scoring system for ICH, we chose to compare the ICHOP scores to it for purposes of evaluating discrimination. The ICHOP₃ and ICHOP₁₂ models provide higher discrimination (AUC) in both the derivation and validation cohorts. However, the difference was only statistically significant in the derivation cohort. There are several possible explanations for this finding. First, a post-hoc power analysis demonstrated that the validation cohort may have been underpowered to demonstrate effect, with the probability of a type II error being greater than 20% in both 3- and 12-month analyses. Second, functional outcomes in the validation cohort were much more favorable overall than in the derivation cohort, a discrepancy that may have affected statistical analysis. In the derivation cohort, 61% (3 months) and 62% (12 months) of patients demonstrated poor functional outcome while only 11% (3 months) and 4% (12 months) demonstrated poor functional outcome in the validation cohort, which is atypical for intracerebral hemorrhage where mortality alone can be greater than 50%. Thus, it is possible that the overall acuity of illness was highly discrepant between the two cohorts, which could have affected the results of our analysis. Prospective validation is the next step in evaluating the utility and universal applicability of the ICHOP scores.

This study has several limitations. First, our derivation cohort and, particularly, our validation cohort were relatively small compared to recent similar studies.^{7,19} Many of our patients were lost to follow up or missing information, both of which contributed to our smaller sample size. However, the derivation cohort sample size was large enough to develop the scoring system and satisfactorily compare it to the ICH Score. Second, the ICHOP₃ score is a fairly complex scoring system, incorporating five variables, including the APACHE II score, which already involves a highly laborious calculation. Although the ICHOP₁₂ score is more digestible, utilizing only four variables, and calculable via a table rather than through a formula, it is still relatively cumbersome. However, given the ubiquitous nature of smartphones and computers, particularly in a hospital setting, creating a smartphone app or a web applet to calculate these scores may offset the complex nature of performing a manual calculation. Finally, both the derivation and validation cohorts were taken from tertiary urban academic medical centers; thus, we have no evidence that this model will prove useful in rural or community environments that serve less complex or more homogenous patient populations.

Conclusions

The ideal scoring system for any disease is easy to calculate, highly discriminatory, and universally applicable. When creating the ICHOP score, we chose to sacrifice ease of calculation for increased discrimination, and preliminary external findings hint that it may be applicable in different settings. Though some physicians treating ICH patients may be hesitant to calculate ICHOP scores, both the high morbidity of the illness and the overwhelming burden that it places upon the healthcare system suggest that even smaller improvements can be beneficial to improving allocation of resources and positively impacting care for patients suffering from such a devastating illness.

Acknowledgments

Funding: This work was supported by the National Institutes of Health [NIA T35AG044303]

Abbreviations

APACHE II	Acute Physiology and Chronic Health Evaluation II
AUC	Area under the receiver operating characteristic curve
CT	Computerized Tomography
CUMC	Columbia University Medical Center
GCS	Glasgow Coma Scale
ICH	Intracerebral hemorrhage
ICHOP	Intracerebral hemorrhage outcomes project
IRB	Institutional Review Board
mRS	modified Rankin Scale
NIHSS	National Institutes of Health Stroke Scale

References

1. Qureshi AI, Mendelow AD, Hanley DF. Intracerebral haemorrhage. *Lancet*. 2009; 373(9675):1632–44. [PubMed: 19427958]
2. Flaherty ML, Haverbusch M, Sekar P, Kissela B, Kleindorfer D, Moomaw CJ, et al. Long-term mortality after intracerebral hemorrhage. *Neurology*. 2006; 66(8):1182–6. [PubMed: 16636234]
3. Taylor TN, Davis PH, Torner JC, Holmes J, Meyer JW, Jacobson MF. Lifetime cost of stroke in the United States. *Stroke*. 1996; 27(9):1459–66. [PubMed: 8784113]
4. Roger VL, Go AS, Lloyd-Jones DM, Benjamin EJ, Berry JD, Borden WB, et al. Heart disease and stroke statistics–2012 update: a report from the American Heart Association. *Circulation*. 2012; 125(1):e2–e220. [PubMed: 22179539]
5. Weimar C, Benemann J, Diener HC. Development and validation of the Essen Intracerebral Haemorrhage Score. *J Neurol Neurosurg Psychiatr*. 2006; 77(5):601–5. [PubMed: 16354736]
6. Hemphill JC, Bonovich DC, Besmertis L, Manley GT, Johnston SC. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke*. 2001; 32(4):891–7. [PubMed: 11283388]
7. Ji R, Shen H, Pan Y, Wang P, Liu G, Wang Y, et al. A novel risk score to predict 1-year functional outcome after intracerebral hemorrhage and comparison with existing scores. *Crit Care*. 2013; 17(6):R275. [PubMed: 24289116]
8. Neidert MC, Lawton MT, Mader M, Seifert B, Valavanis A, Regli L, et al. The AVICH Score: A Novel Grading System to Predict Clinical Outcome in Arteriovenous Malformation-Related Intracerebral Hemorrhage. *World Neurosurg*. 2016; 92:292–7. [PubMed: 27150647]
9. Rost NS, Smith EE, Chang Y, Snider RW, Chanderraj R, Schwab K, et al. Prediction of functional outcome in patients with primary intracerebral hemorrhage: the FUNC score. *Stroke*. 2008; 39(8):2304–9. [PubMed: 18556582]
10. Appelboom G, Bruce SS, Han J, Piazza M, Hwang B, Hickman ZL, et al. Functional outcome prediction following intracerebral hemorrhage. *J Clin Neurosci*. 2012; 19(6):795–8. [PubMed: 22516544]
11. Bruce SS, Appelboom G, Piazza M, Hwang BY, Kellner C, Carpenter AM, et al. A comparative evaluation of existing grading scales in intracerebral hemorrhage. *Neurocrit Care*. 2011; 15(3):498–505. [PubMed: 21394545]
12. Gilmore EJ, Sheth KN. Improving ascertainment and communication of prognosis in intracerebral hemorrhage. *Crit Care*. 2013; 17(6):1021. [PubMed: 24321182]

13. Heeley E, Anderson CS, Woodward M, Arima H, Robinson T, Stapf C, et al. Poor utility of grading scales in acute intracerebral hemorrhage: results from the INTERACT2 trial. *Int J Stroke*. 2015; 10(7):1101–7. [PubMed: 26044442]
14. Hemphill JC, Farrant M, Neill TA. Prospective validation of the ICH Score for 12-month functional outcome. *Neurology*. 2009; 73(14):1088–94. [PubMed: 19726752]
15. Huang Y, Chen J, Zhong S, Yuan J. Role of APACHE II scoring system in the prediction of severity and outcome of acute intracerebral hemorrhage. *Int J Neurosci*. 2016; 126(11):1020–4. [PubMed: 26393395]
16. Weimar C, Roth M, Willig V, Kostopoulos P, Benemann J, Diener HC. Development and validation of a prognostic model to predict recovery following intracerebral hemorrhage. *J Neurol*. 2006; 253(6):788–93. [PubMed: 16525882]
17. Woo D, Broderick JP, Kothari RU, Lu M, Brott T, Lyden PD, et al. Does the National Institutes of Health Stroke Scale favor left hemisphere strokes? NINDS t-PA Stroke Study Group. *Stroke*. 1999; 30(11):2355–9. [PubMed: 10548670]
18. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988; 19(5):604–7. [PubMed: 3363593]
19. Witsch J, Frey HP, Patel S, Park S, Lahiri S, Schmidt JM, et al. Prognostication of long-term outcomes after subarachnoid hemorrhage: The FRESH score. *Ann Neurol*. 2016; 80(1):46–58. [PubMed: 27129898]
20. Claassen J, Vu A, Kreiter KT, Kowalski RG, Du EY, Ostapkovich N, et al. Effect of acute physiologic derangements on outcome after subarachnoid hemorrhage. *Crit Care Med*. 2004; 32(3):832–8. [PubMed: 15090970]
21. LoPresti MA, Bruce SS, Camacho E, Kunchala S, Dubois BG, Bruce E, et al. Hematoma volume as the major determinant of outcomes after intracerebral hemorrhage. *J Neurol Sci*. 2014; 345(1–2):3–7. [PubMed: 25034055]
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*. 1988; 44(3):837–845. [PubMed: 3203132]

Highlights

- A novel prognostication system for functional outcomes following ICH is proposed.
- Scores incorporate systemic physiological data and pre-morbid functional status.
- Novel discrimination is higher than original ICH Score for prognosticating outcome.
- Preliminary external validation suggests applicability to different populations.

Figure 1.

Percentage of patients at each ICHOP₃ (A) and ICHOP₁₂ (B) level with poor functional outcome at 3 or 12 months, respectively.

Figure 2.

Receiver operating characteristic curves for ICHOP₃ (A) and ICHOP₁₂ (B) scores. The area under the curve (AUC) for ICHOP₃ was 0.89 with a 95% confidence interval of 0.858–0.926. The AUC for ICHOP₁₂ was 0.89 with a 95% confidence interval of 0.836–0.910.

Table 1

Baseline characteristics of ICH patients in 3-month derivation cohort (n=365). Comparison to full database (n=575) is shown, with number of patients with available data noted in parentheses.

Characteristic (median $\pm$ SD unless noted)	Derivation Cohort	Full database
Age	64 $\pm$ 17.8	64 $\pm$ 17.7 (n=575)
GCS	10 $\pm$ 4.5	11 $\pm$ 4.5 (n=566)
NIHSS	13 $\pm$ 12.3	11 $\pm$ 12.15 (n=539)
APACHE II	15 $\pm$ 8.4	14 $\pm$ 8.4 (n=534)
APACHE II Physiologic (without GCS component)	10 $\pm$ 5.6	9 $\pm$ 5.5 (n=532)
ICH Score	2 $\pm$ 1.3	2 $\pm$ 1.3 (n=536)
Hematoma volume	14.2 $\pm$ 32.6	13.3 $\pm$ 33.9 (n=559)
Intraventricular expansion (% of patients)	54.4%	53.0% (n=533)
On anticoagulants (% of patients)	13.0%	13.2% (n=575)
Premorbid mRS	0 $\pm$ 1.5	1 $\pm$ 1.4 (n=548)
mRS at 3 months	4 $\pm$ 1.9	4 $\pm$ 2.0 (n=444)
mRS at 12 months	5 $\pm$ 2.0	4 $\pm$ 2.0 (n=395)
Poor functional outcome at 3 months	61%	62% (n=444)
Poor functional outcome at 12 months	62%	64% (n=395)
Mortality at 3 months	37%	39% (n=444)
Mortality at 12 months	50%	47% (n=395)

Table 2

Categorical groupings of GCS, NIHSS, and hematoma volume with descriptive designations. Values assigned to each category for use in the ICHOP Scores are shown at the right of each box.

Component	Value	Descriptor	ICHOP ₃
GCS	3–8	Severe	2
	9–12	Moderate	1
	13–15	Minor	0
NIHSS	21–42	Highly Severe	4
	16–20	Moderately Severe	3
	5–15	Moderate	2
	1–4	Minimal	1
	0	No symptoms	0
Hematoma Volume	30+ mL	High	2
	1–30 mL	Low	1
	0 mL	None	0

Table 3

Calculation of the ICHOP₁₂ score based on values ascribed to grouped factor scores.

Component	Value	ICHOP ₁₂ Score Points
GCS	3–8	+6
	9–12	+3
	13–15	+0
NIHSS	21–42	+4
	16–20	+3
	5–15	+2
	1–4	+1
	0	+0
APACHE II _{phys}	0–56	+ (APACHE II _{phys}) × 2
Premorbid mRS	5	+20
	4	+16
	3	+12
	2	+8
	1	+4
	0	+0
ICHOP ₁₂ Total Score		0–142

Table 4

Performance measures for ICHOP₃ and ICHOP₁₂ scores alongside ICH Scores for predicting functional outcome at 3 and 12 months in the derivation cohort.

	ICHOP ₃	ICH (3 month)	ICHOP ₁₂	ICH (12 month)
AUC (95% CI)	0.89 (0.858–0.926)	0.84 (0.806–0.874)	0.87 (0.836–0.910)	0.81 (0.769–0.854)
McFadden R ²	0.40	0.29	0.36	0.24
Cox/Snell R ²	0.43	0.34	0.39	0.28
Nagelkerke R ²	0.57	0.45	0.52	0.37