# Clinician judgment vs formal scales for predicting intracerebral hemorrhage outcomes

David Y. Hwang, MD
Cameron A. Dell, BS
Mary J. Sparks, RN, BSN
Tiffany D. Watson
Carl D. Langefeld, PhD
Mary E. Comeau, MA
Jonathan Rosand, MD, MSc
Thomas W.K. Battey, BS
Sebastian Koch, MD
Mario L. Perez, MD
Michael L. James, MD
Jessica McFarlin, MD
Jennifer L. Osborne, RN, BSN
Daniel Woo, MD
Steven J. Kittner, MD, MPH
Kevin N. Sheth, MD

Correspondence to Dr. Hwang: david.hwang@yale.edu

## ABSTRACT

**Objective:** To compare the performance of formal prognostic instruments vs subjective clinical judgment with regards to predicting functional outcome in patients with spontaneous intracerebral hemorrhage (ICH).

**Methods:** This prospective observational study enrolled 121 ICH patients hospitalized at 5 US tertiary care centers. Within 24 hours of each patient's admission to the hospital, one physician and one nurse on each patient's clinical team were each asked to predict the patient's modified Rankin Scale (mRS) score at 3 months and to indicate whether he or she would recommend comfort measures. The admission ICH score and FUNC score, 2 prognostic scales selected for their common use in neurologic practice, were calculated for each patient. Spearman rank correlation coefficients ($r$) with respect to patients' actual 3-month mRS for the physician and nursing predictions were compared against the same correlation coefficients for the ICH score and FUNC score.

**Results:** The absolute value of the correlation coefficient for physician predictions with respect to actual outcome (0.75) was higher than that of either the ICH score (0.62, $p = 0.057$) or the FUNC score (0.56, $p = 0.01$). The nursing predictions of outcome ($r = 0.72$) also trended towards an accuracy advantage over the ICH score ($p = 0.09$) and FUNC score ($p = 0.03$). In an analysis that excluded patients for whom comfort care was recommended, the 65 available attending physician predictions retained greater accuracy ($r = 0.73$) than either the ICH score ($r = 0.50$, $p = 0.02$) or the FUNC score ($r = 0.42$, $p = 0.004$).

**Conclusions:** Early subjective clinical judgment of physicians correlates more closely with 3-month outcome after ICH than prognostic scales. *Neurology*® 2016;86:126–133

## GLOSSARY

**ERICH** = Ethnic/Racial Variations of Intracerebral Hemorrhage; **GCS** = Glasgow Coma Scale; **ICH** = intracerebral hemorrhage; **mRS** = modified Rankin Scale; **NICU** = neuroscience intensive care unit.

Spontaneous intracerebral hemorrhage (ICH) is a devastating stroke subtype, with high rates of mortality and long-term disability.[1–3] Early impressions of ICH prognosis are particularly important, as decisions regarding life-sustaining therapy often occur within the first days of hospitalization and carry a self-fulfilling prophecy risk.[4–8]

Numerous formal clinical grading scales for patients with spontaneous ICH estimate long-term mortality or functional outcome, using clinical and radiographic variables from admission.[9–16] While some of these outcome calculators have become commonplace, with inclusion in the documentation quality metrics among US Comprehensive Stroke Centers, few have been validated independently and prospectively.[17–19] Even fewer have been compared to the early clinical judgment of bedside clinicians with regards to their accuracy.[20]

The present study sought to compare 2 of the most commonly used ICH clinical scales in neurologic practice, the ICH score and the FUNC score, to the early subjective clinical

Supplemental data at Neurology.org

From the Division of Neurocritical Care and Emergency Neurology (D.Y.H., K.N.S.), Department of Neurology, Yale School of Medicine, New Haven, CT; the Maryland Stroke Center (C.A.D., M.J.S., T.D.W.), Baltimore; the Center for Public Health Genomics and Department of Biostatistical Sciences (C.D.L., M.E.C.), Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC; the Center for Human Genetic Research (J.R., T.W.K.B.), Boston, MA; the University of Miami (S.K., M.L.P.), Miller School of Medicine, FL; Duke University Medical Center (M.L.J., J.M.), Durham, NC; the Department of Neurology (J.L.O., D.W.), University of Cincinnati College of Medicine, OH; and the Baltimore Veterans Administration Medical Center and University of Maryland School of Medicine (S.J.K.).

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

judgment of physicians and nurses with regards to predicting functional outcome in patients with spontaneous ICH.[9,14] The prespecified hypothesis was that the 2 formal scales would more accurately predict modified Rankin Scale (mRS) score at 3 months after ICH compared to early subjective predictions of bedside physicians and nurses, as measured by each group's correlation with actual outcome.

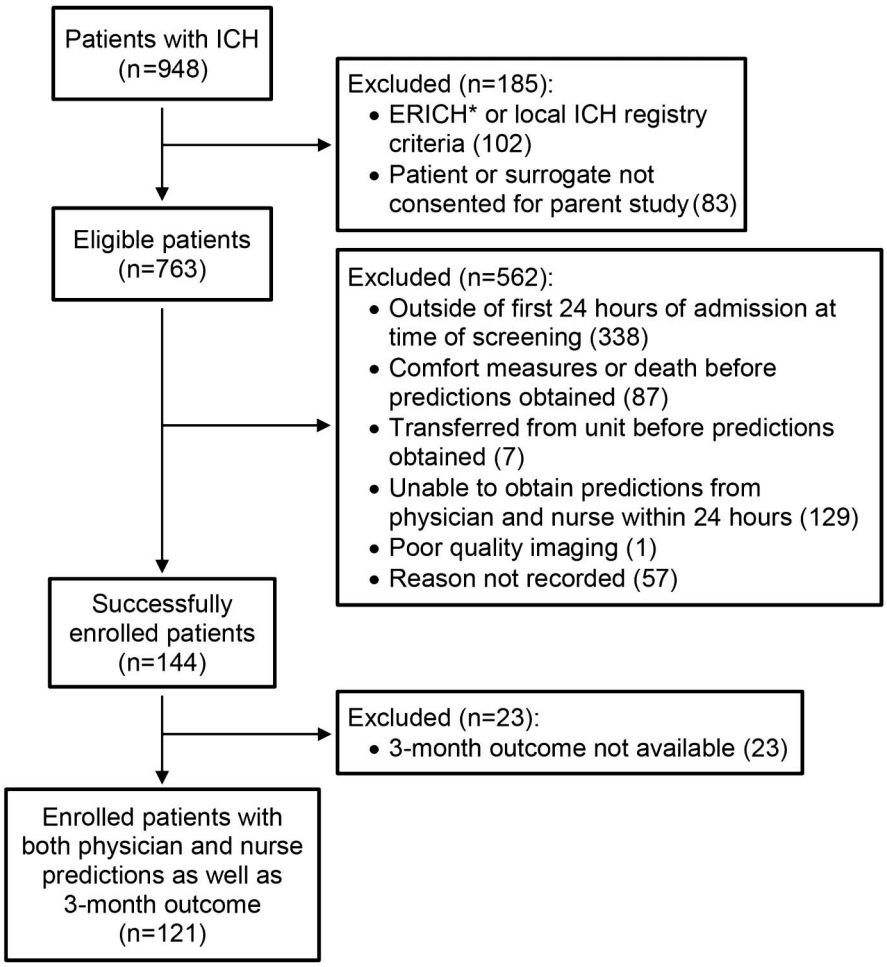
**METHODS** **Study design and setting.** This prospective multicenter observational study enrolled patients with spontaneous ICH at 5 neuroscience intensive care units (NICUs) within 5 academic medical centers in the United States from April 1, 2011, to December 31, 2013. Within 24 hours of each patient's admission to the hospital, one physician and one nurse on the primary team were asked to predict the patient's functional outcome at 3 months. This early 24-hour period for collecting predictions was incorporated into the study design so that predictions were made before actual decisions regarding aggressiveness of care for severely affected patients dictated their outcomes. The values for the ICH score and FUNC score were then calculated for each patient using admission clinical and radiographic variables. The actual 3-month outcome for each patient was obtained by personnel blinded to the original predictions to enable comparison of the predictive accuracy of providers against formal scales.

**Standard protocol approvals, registrations, and patient consents.** This study was approved by the institutional ethical standards committee on human experimentation at all 5 sites. All patient and clinician participants gave informed consent (or via surrogate for incapacitated patients) before taking part.

**Participants.** ***Patients.*** An eligibility criterion for this study was enrollment within 24 hours of admission. Other eligibility criteria were from the ongoing Ethnic/Racial Variations of Intracerebral Hemorrhage (ERICH) Study, a large prospective observational study at 19 US centers examining genetic and epidemiologic risk factors and outcomes for ICH[21]:

1. Age 18 years or greater.
2. Resident for at least 6 months within 75 miles of recruiting center.
3. Diagnosis of spontaneous ICH, including warfarin-associated ICH.
4. Ability of the patient or legal representative to provide informed consent.

**Figure** **Flow chart of enrollment of patients with intracerebral hemorrhage (ICH) and their clinicians making predictions**




*Ethnic/Racial Variations of Intracerebral Hemorrhage study protocol.

**Table 1** Characteristics of patients with ICH

| Characteristics | Total (n = 121) |
|---|---|
| Mean (SD) age, y | 66.8 (14.5) |
| Median age, y | 67 |
| Women | 55 (45.5) |
| Race | |
| White | 67 (55.4) |
| Black | 50 (41.3) |
| Other/unknown | 4 (3.3) |
| Hispanic ethnicity | 16 (13.2) |
| Pre-ICH comorbidities | |
| Cognitive impairment | 20 (16.5) |
| Hypertension | 99 (81.8) |
| Diabetes mellitus | 37 (30.6) |
| Coronary artery disease | 23 (19.0) |
| Atrial fibrillation | 17 (14.0) |
| Ischemic stroke | 20 (16.5) |
| ICH (history) | 11 (9.1) |
| Malignancy | 14 (11.6) |
| Alcohol (2 or more drinks per day on average) | 14 (11.6) |
| Smoking | |
| Current | 20 (16.5) |
| Former | 35 (28.9) |
| Cocaine | 7 (5.8) |
| Initial clinical presentation | |
| Glasgow Coma Scale score | |
| 13–15 | 75 (62.0) |
| 5–12 | 38 (31.4) |
| 3–4 | 8 (6.6) |
| Recent warfarin use (within the past 2 weeks) | 15 (12.4) |
| Initial radiographic presentation | |
| ICH volume, mL | |
| <30 | 81 (66.9) |
| 30 ≤ volume ≤ 60 | 20 (16.5) |
| >60 | 20 (16.5) |
| ICH location | |
| Lobar | 46 (38.0) |
| Deep | 62 (51.2) |
| Infratentorial | 13 (10.7) |
| Presence of intraventricular hemorrhage | 55 (45.5) |
| Clinical course | |
| Surgical evacuation | 7 (5.8) |
| Ventriculostomy placement | 28 (23.1) |
| Made CMO | 31 (25.6) |
| Mean (SD) days from admission to CMO status | 6.4 (9.4) |
| Died during hospitalization | 29 (24.0) |
| Site of study enrollment | |
| Site 1 | 49 (40.5) |
| Site 2 | 34 (28.1) |
| Site 3 | 22 (18.2) |
| Site 4 | 12 (9.9) |
| Site 5 | 4 (3.3) |

Abbreviations: CMO = comfort measures only; ICH = intracerebral hemorrhage.
Figures are n (%) of patients unless stated otherwise.

Patients with malignancies that lead to coagulopathy, venous sinus thrombosis, vascular malformations, aneurysms, tumors, hemorrhagic conversion of recent ischemic stroke, or traumatic hemorrhage were excluded.

Each participating center was an ongoing ERICH recruitment center or had a preexisting local prospective ICH registry that enrolled patients with the same eligibility criteria as those listed for ERICH. Patients needed to be enrolled in an existing parent study in order to be considered for this study. Each center screened potential patients for parent studies through review of hospital admission, emergency room, and NICU admission logs.[21] Each patient successfully consented for a parent study was then subsequently screened for this study by assessing whether he or she was still within the first 24 hours of hospital admission.

***Clinicians.*** Within 24 hours of each eligible patient's admission, a single physician and a single nurse on each patient's primary inpatient team were enrolled in the subjective prediction portion of the study by a research coordinator. When available, the attending physician for an eligible patient was enrolled. If the attending was not available, a trainee physician (e.g., fellow, resident) was enrolled, as long as he or she was a member of the patient's treatment team. For each center, informed consent forms were signed in advance by as many members of the physician and nursing staff as possible before the recruitment of patients began, with signed consent by clinicians then verified by research coordinators at the time of each patient's enrollment.

**Variables.** ***Outcome.*** The primary outcome variable obtained for each enrolled patient was the mRS, scored at 3 months.[22]

***Variables predictive of outcome.*** The collected predictors of outcome obtained for each patient were (1) a single physician prediction of 3-month mRS and (2) a single nursing prediction of 3-month mRS. All predictions were obtained within 24 hours of patient admission. As each participating clinician was a member of an enrolled patient's primary team, physicians and nurses were allowed to use any information available to them for formulating prognosis. Concurrently, each physician and nurse was asked whether he or she would recommend comfort care for each enrolled patient.

The calculated predictors of outcome were the ICH score and the FUNC score. Details of how the ICH score and FUNC score are calculated are provided in table e-1 on the *Neurology*® Web site at Neurology.org. The ICH score was selected for this study because it is by far the most validated ICH prognostication tool available.[9] It outputs an ordinal value from 0 to 6, with increasing values correlating with increasing probabilities of poor outcome as measured by

the mRS. The ICH score has been validated for predicting the probabilities of achieving each mRS level by patients at 3, 6, 9, and 12 months.[17] The FUNC score outputs values from 0 to 11; in contrast to the ICH score and the mRS, increasing values correlate with increasing probabilities of favorable outcome.[14] Designed to predict probability of functional independence at 3 months, the FUNC score was selected for this study because of its focus on functional outcome and its initial validation in a survivors-only cohort. The correlations of the ICH score and FUNC score with patient outcome were derived using patient data taken from time of initial hospital evaluation.

**Data sources and measurement.** All clinical variables for patients were obtained via direct chart abstraction at enrolling centers, with Glasgow Coma Scale (GCS) taken from time of initial evaluation. All radiographic variables were consistently assessed from initial head CT scans. For patients enrolled in ERICH, baseline CT images were sent to the ERICH neuroimaging core site (Massachusetts General Hospital) for independent assessment of radiographic variables.[21] Patients not enrolled in ERICH had their initial imaging assessed locally by a blinded reviewer.

Each clinician indicated his or her best estimate of the mRS a patient would achieve at 3 months after ICH via paper survey. This survey included written detailed descriptions of each mRS value. Whether the respondent would recommend comfort care for the patient was also explicitly asked on the survey. Local research coordinators entered all predictions and clinician variables into a centralized password-protected online database for analysis.

Actual 3-month mRS for each enrolled patient was obtained via telephone interview by blinded ERICH staff or by research coordinators responsible for local site ICH registries where applicable.

**Potential bias.** As active members of the enrolled patients' treatment teams, physicians and nurses enrolled in this study were not blinded to clinical information outside of the variables already incorporated into the ICH score and FUNC score. This design ensured that the predictive accuracy of observed clinicians mirrored that of real-life situations for generalizability of study results to clinical practice.

**Statistical methods and study size.** Patient and clinician variables were described by calculating counts, percentages, and means ± SDs and medians, as appropriate. To describe clinician prognostication data, matched contingency tables of predicted mRS vs actual 3-month mRS were generated separately for physicians and nurses. Weighted kappa values were computed as a measure of agreement beyond chance. To test whether there were directional trends (either optimistic or pessimistic) for physician and nursing predictions, a $p$ value test of symmetry for each contingency table was also computed.

The accuracies of physician and nurse predictions and of the ICH score and FUNC score were quantified by computing the Spearman rank correlation coefficient for the output of each group with actual 3-month outcome. The accuracy of attending physicians and trainees were also assessed separately using this method. The Wald test was computed to test whether there were significant differences in the value of the correlation coefficient between groups.

Separate prespecified analyses using these same statistical methods were performed on cohorts of only those patients surviving to 3 months and only those patients for whom neither the physician nor nurse surveyed indicated that he or she would initially recommend comfort care. These cohorts were designed to minimize bias from clinicians' early knowledge of inevitable limitations of care. The accuracy of physicians and nurses compared to the ICH score and FUNC score was also assessed using the same statistical method within different subgroups of patients divided by markers of disease severity, i.e., admission GCS score or initial ICH volume, in exploratory analyses that were not prespecified.

Although the sample sizes reported here represent an unprecedented effort to explore these questions, there remain limitations relative to robustly delineating plausible health care provider or patient characteristics that might explain patterns observed. The sample sizes are sufficient, however, for exploring large differences among groups.

**Table 2** Characteristics of physicians and nurses making outcome predictions

| Characteristics | Total (n = 36) |
|---|---|
| Attending physicians | |
| Female | 4 (13.5) |
| Mean (SD) age, y | 41.1 (8.2) |
| Primary neurologic subspecialty | 24 (66.7) |
| Mean (SD) years in practice | 8.8 (7.8) |
| Trainee physicians | n = 16 |
| Female | 2 (12.5) |
| Mean (SD) age, y | 32.5 (4.4) |
| Primary neurologic subspecialty | 14 (87.5) |
| Mean (SD) years in practice | 3.9 (2.4) |
| Nurses | n = 76 |
| Female | 66 (86.8) |
| Mean (SD) age, y | 33.7 (10.2) |
| Education | |
| Associate's degree | 10 (13.2) |
| Bachelor's degree | 59 (77.6) |
| Master's degree | 7 (9.2) |
| Primary specialty | |
| Neuroscience | 55 (72.4) |
| Critical care | 19 (25.0) |
| Official certification | |
| Neuroscience | 9 (11.8) |
| Critical care | 6 (7.9) |
| Mean (SD) years in practice | 7.3 (8.1) |
| Mean (SD) years of neuroscience intensive care unit experience | 4.8 (6.8) |

Figures are n (%) of care providers unless stated otherwise.

## RESULTS

**Enrollment.** The figure contains details regarding the enrollment from April 2011 to December 2013 of 121 ICH patients, with both early physician and nurse predictions as well as 3-month outcomes.

**Patients.** Table 1 shows the baseline characteristics of the 121 patients.

**Clinicians.** Table 2 shows characteristics of the physicians and nurses making outcome predictions. Thirty-six attending physicians and 16 trainee physicians made 121

Table 3 Correlation with 3-month outcome of clinician predictions compared to ICH outcome scales

| | | ICH score | | FUNC score | |
|---|---|---|---|---|---|
| Clinician group | $r^a$ | $r^a$ | $p$ | $r^b$ | $p$ |
| **Total patient cohort (n = 121)** | | | | | |
| All physicians | 0.75 | 0.62 | 0.057 | −0.56 | 0.009 |
| Attendings (n = 91) | 0.78 | | 0.02 | | 0.003 |
| Trainees (n = 30) | 0.64 | | 0.87 | | 0.56 |
| Nurses | 0.72 | | 0.16 | | 0.03 |
| **For survivors at 3 mo (n = 78)** | | | | | |
| All physicians | 0.62 | 0.34 | 0.02 | −0.33 | 0.02 |
| Attendings (n = 59) | 0.72 | | 0.002 | | 0.001 |
| Trainees (n = 19) | 0.21 | | 0.61 | | 0.64 |
| Nurses | 0.58 | | 0.06 | | 0.05 |
| **Without comfort care bias[c] (n = 89)** | | | | | |
| All physicians | 0.68 | 0.50 | 0.07 | −0.42 | 0.01 |
| Attendings (n = 65) | 0.73 | | 0.02 | | 0.004 |
| Trainees (n = 24) | 0.52 | | 0.91 | | 0.60 |
| Nurses | 0.64 | | 0.17 | | 0.04 |

Abbreviation: ICH = intracerebral hemorrhage.
[a] $r$ represents the Spearman rank correlation coefficient vs 3-month actual outcome as measured by the modified Rankin Scale. A perfect coefficient with an absolute value of 1 occurs when the 2 tested variables are a perfect monotonic function of the other. The listed $p$ values represent the probabilities that differences in $r$ for the ICH score and FUNC score with respect to those for the clinician groups are due to chance.
[b] Because a high FUNC score value is ideally supposed to correlate with an eventual low modified Rankin Scale value, the Spearman rank correlation coefficients calculated for FUNC scores are negative. The absolute value can be compared with the other calculated $r$ values.
[c] Patients for whom either a physician or a nurse indicated on the prediction form that he or she would be inclined to recommend comfort care to the family were excluded from this analysis.

predictions, while the 121 nursing predictions were made by 76 nurses.

**Clinician predictions.** A table of predicted mRS vs actual 3-month mRS for physicians and nurses, respectively, was created (table e-2). The weighted kappa measure of agreement suggested a moderate improvement over chance for predictions of actual mRS score, with the nursing, physician (attendings and trainees), and attending-physician-only weighted kappas, asymptotic standard error, and 95% confidence intervals being 0.47 ± 0.05 (0.39–0.56), 0.49 ± 0.04 (0.41–0.57), and 0.53 ± 0.05 (0.44–0.63), respectively.

The observed 7 × 7 tables for both physicians and nurses showed significant departures from symmetry ($p = 0.03$ and $p = 0.04$, respectively). Careful comparison of the frequency of actual mRS outcomes with the frequency of predicted mRS outcomes reveals a tendency towards optimistic errors for both physicians and nurses when making early outcome predictions. While 46 patients (37.1%) experienced 3-month mortality, the physician and nursing group each only predicted 18 patients (14.5%) to be dead at follow-up.

**Clinician predictions vs ICH score and FUNC score.** Table 3 shows the Spearman rank correlation coefficients ($r$) of the physician predictions and nursing predictions compared to those of the ICH score and FUNC score, relative to the actual 3-month mRS, including separate analyses of a cohort of all patients who were alive at 3 months and the cohort of patients after those were excluded for whom either a physician or nurse predicted that he or she would be inclined to recommend comfort care to the patient's family.

To explore the possibility that the observed predictive advantage of clinicians over formal scales may simply be due to the ability of physicians and nurses to prognosticate patients who present very well (e.g., high GCS, small ICH volume) or very poorly (e.g., low GCS, large ICH volume) with a high degree of accuracy, the comparative analyses displayed in table 4 were conducted.

First, the entire sample of patients was divided into 2 subgroups by GCS, with very low (i.e., 3–4) or high (i.e., 13–15) GCS scores in one group (n = 83) and middle range (5–12) GCS scores in another (n = 38). Second, the entire sample of patients was divided into 2 subgroups by ICH volume, with small

**Table 4** Correlation with 3-month outcome of clinician predictions vs ICH outcome scales for notable patient subgroups

| | | ICH score | | FUNC score | |
|---|---|---|---|---|---|
| Clinician group, divided by patient subgroup | $r$[a] | $r$[a] | $p$ | $r$[b] | $p$ |
| Glasgow Coma Scale | | | | | |
| Middle range: 5–12 (n = 38) | | | | | |
| Physicians | 0.59 | 0.14 | 0.02 | −0.37 | 0.23 |
| Nurses | 0.55 | | 0.05 | | 0.34 |
| Very low or high: 3–4 or 13–15 (n = 83) | | | | | |
| Physicians | 0.69 | 0.55 | 0.15 | −0.52 | 0.09 |
| Nurses | 0.70 | | 0.12 | | 0.07 |
| ICH volume | | | | | |
| Moderate: between 30 and 60 mL (n = 20) | | | | | |
| Physicians | 0.60 | 0.33 | 0.31 | −0.01 | 0.05 |
| Nurses | 0.64 | | 0.23 | | 0.03 |
| Small or large: <30 or >60 mL (n = 101) | | | | | |
| Physicians | 0.75 | 0.62 | 0.08 | −0.56 | 0.02 |
| Nurses | 0.71 | | 0.25 | | 0.08 |

Abbreviation: ICH = intracerebral hemorrhage.
[a] $r$ represents the Spearman rank correlation coefficient vs 3-month actual outcome as measured by the modified Rankin Scale. A perfect coefficient with an absolute value of 1 occurs when the 2 tested variables are a perfect monotonic function of the other. The listed $p$ values represent the probabilities that differences in $r$ for the ICH score and FUNC score with respect to those for the clinician groups are due to chance.
[b] Because a high FUNC score value is ideally supposed to correlate with an eventual low modified Rankin Scale value, the Spearman rank correlation coefficients calculated for FUNC scores are negative. The absolute value can be compared with the other calculated $r$ values.

(<30 mL) and large (>60 mL) hemorrhages in one group (n = 101) and moderate sized hemorrhages (30 mL ≤ volume ≤ 60 mL) in another (n = 20). While these analyses were not prespecified and are limited (in the situation of moderate volumes) by low numbers, the outcome predictions of both physicians and nurses still trended towards greater predictive accuracy than either the ICH score or FUNC score for both GCS subgroups and both ICH volume subgroups alike.

**DISCUSSION** Contrary to the hypothesis, for patients with spontaneous ICH, the early subjective clinical judgment of physicians within 24 hours of patient admission correlates more closely with 3-month outcome than 2 of the most widely used ICH outcome prognostication tools, the ICH score and the FUNC score. For attending physicians, this finding is true even when considering only those patients for whom the primary care team is not considering early institution of comfort measures. The accuracy of physicians over the 2 formal scales may be present not only when patients' initial clinical presentations are clearly favorable (e.g., high GCS, small ICH volume) or unfavorable (e.g., low GCS, large ICH volume) but even when clinical presentations are more ambiguous in nature. Nursing predictions of outcome trend towards similar advantages over the ICH score and FUNC score as well. When subjective prognostic errors are made for mortality, both physicians and nurses tend towards optimism.

Two particular strengths of our multicenter study are that all physician and nursing predictions of outcome were obtained within 24 hours of patient admission and both a survivors-only cohort and a cohort of patients without a comfort care bias were included. With these key study aspects in mind, practical challenges and study limitations should be acknowledged. First, a significant number of patients enrolled in local parent ICH studies did not undergo physician and nursing predictions within 24 hours. The percentage of patients in our study with a GCS above 8 (81.0%) is slightly higher than the percentages seen in other ICH cohorts published from the United States; however, the percentages of patients with respect to other relevant demographic and clinical factors (e.g., patient gender, age, ICH location, ICH volume) are comparable to past cohorts (table e-3).[9,14,17,19] Overall mortality in this cohort mirrors that seen in prior publications.[3,19] Second, this study does not include specific data on clinicoradiographic deterioration of enrolled patients

within the first 24 hours or information on ICH location outside of characterizations incorporated into the clinical scales (i.e., lobar, deep, infratentorial). It is possible that having such information might provide further insight into the reasons driving the superiority of clinician predictions. Third, clinicians making predictions were not formally trained in mRS scoring, although the paper survey that was distributed did contain the details of the mRS as a reference.

While comparisons of clinical judgment against outcome prediction tools in general critical care populations have been studied,[23] our novel study prospectively compares the accuracy of commonly used formal prognostication scales against subjective clinician judgment specifically for ICH patients. Despite their importance, prognostic studies of subjective clinician judgments for critical care neurology have been few.[24,25] One recent study in ischemic stroke compared prognostic accuracy of physician judgment to a widely available outcome prediction tool and concluded that the formal tool was superior.[26] However, physicians were given hypothetical case vignettes to prognosticate. Prognostication of stroke, whether ischemic or hemorrhagic, in a true clinical setting may allow clinicians to incorporate a range of decision-making factors not provided in hypothetical vignettes.

While the authors of the ICH score and FUNC score have already warned against using their instruments for the prognostication of individual patient outcome,[14,17] the findings of this study further emphasize that there is currently no substitute for experienced clinician judgment. The profile of clinicians in this multicenter study suggests that the results are generalizable at least to academic medical centers with neurologic specialists, with the accuracy of clinicians at community hospitals remaining an open question. While avoidance of the self-fulfilling prophesy is noble in medicine, the fact that both physicians and nurses tend to err on the side of optimism for prognosticating ICH mortality suggests that the field of neurocritical care in general may benefit from being guarded when talking with families about outcome for these patients, in particular when discussions involve potentially committing patients to interventions such as tracheotomy or percutaneous endoscopic gastrostomy. The prognosis from the physician team alone is not the only factor that surrogate decision-makers consider in such situations but certainly influences ultimate decision-making.[27]

The subjective factors that physicians and nurses incorporate into their determination of ICH prognosis, including psychosocial aspects of care, and the reasons for general prognostic optimism, especially regarding patient survival, are unanswered questions open for future study. Gaining a better understanding of particular clinical situations in which tools such as the ICH score and FUNC score may or may not be more accurate than clinician judgment, especially with regards to patients who present with moderate-sized hemorrhages, would yield great improvement in the inherently uncertain process of prognostication.

## AUTHOR CONTRIBUTIONS

All authors contributed to the study design and data analysis and approved the final version of the manuscript. D.Y.H. monitored data collection for all sites; cleaned, analyzed, and interpreted the data; and drafted the paper. C.A.D., M.J.S., and T.D.W. recruited subjects and collected data for their local site and monitored data collection for all sites. C.D.L. and M.E.C. performed statistical analysis and contributed to study conception. T.W.K.B., M.L.P., J.M., and J.L.O. recruited subjects and collected data at respective sites. J.R., S.K., M.L.J., D.W., S.J.K., and K.N.S. contributed to study conception and supervised data collection at respective sites. K.N.S. initiated the study and designed data collection tools. D.W. and K.N.S. obtained funding for the study.


## ACKNOWLEDGMENT

The authors thank Christian Ramos, Division of Neurocritical Care and Emergency Neurology, Department of Neurology, Yale School of Medicine, for assistance with cleaning the data. The Wake Forest School of Medicine Center for Public Health Genomics provided computational support.

## STUDY FUNDING

Supported by the American Heart Association (Clinical Research Training Award 11CRP5480009 to K.N.S.) and National Institute of Neurological Disorders and Stroke (R01NS059727 to J.R. and U01NS069763 to D.W.). The sources of funding had no role in study design, data collection, analyses, interpretations, or decision to submit the article for publication.


## DISCLOSURE

The authors report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.


*Received May 22, 2015. Accepted in final form September 3, 2015.*


## REFERENCES


1. Go AS, Mozaffarian D, Roger VL, et al. Heart disease and stroke statistics: 2014 update: a report from the American Heart Association. Circulation 2014;129:e28–e292.
2. Zahuranec DB, Lisabeth LD, Sanchez BN, et al. Intracerebral hemorrhage mortality is not changing despite declining incidence. Neurology 2014;82:2180–2186.
3. van Asch CJ, Luitse MJ, Rinkel GJ, van der Tweel I, Algra A, Klijn CJ. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. Lancet Neurol 2010;9:167–176.
4. Becker KJ, Baxter AB, Cohen WA, et al. Withdrawal of support in intracerebral hemorrhage may lead to self-fulfilling prophecies. Neurology 2001;56:766–772.
5. Hemphill JC III, Newman J, Zhao S, Johnston SC. Hospital usage of early do-not-resuscitate orders and outcome after intracerebral hemorrhage. Stroke 2004;35:1130–1134.
6. Zahuranec DB, Brown DL, Lisabeth LD, et al. Early care limitations independently predict mortality after intracerebral hemorrhage. Neurology 2007;68:1651–1657.
7. Holloway RG, Gramling R, Kelly AG. Estimating and communicating prognosis in advanced neurologic disease. Neurology 2013;80:764–772.

8. Morgenstern LB, Hemphill JC III, Anderson C, et al. Guidelines for the management of spontaneous intracerebral hemorrhage: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2010;41:2108–2129.
9. Hemphill JC III, Bonovich DC, Besmertis L, Manley GT, Johnston SC. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. Stroke 2001;32:891–897.
10. Cheung RT, Zou LY. Use of the original, modified, or new intracerebral hemorrhage score to predict mortality and morbidity after intracerebral hemorrhage. Stroke 2003;34:1717–1722.
11. Godoy DA, Pinero G, Di Napoli M. Predicting mortality in spontaneous intracerebral hemorrhage: can modification to original score improve the prediction? Stroke 2006;37:1038–1044.
12. Weimar C, Benemann J, Diener HC; German Stroke Study Collaboration. Development and validation of the Essen Intracerebral Haemorrhage Score. J Neurol Neurosurg Psychiatry 2006;77:601–605.
13. Ruiz-Sandoval JL, Chiquete E, Romero-Vargas S, Padilla-Martinez JJ, Gonzalez-Cornejo S. Grading scale for prediction of outcome in primary intracerebral hemorrhages. Stroke 2007;38:1641–1644.
14. Rost NS, Smith EE, Chang Y, et al. Prediction of functional outcome in patients with primary intracerebral hemorrhage: the FUNC score. Stroke 2008;39:2304–2309.
15. Cho DY, Chen CC, Lee WY, Lee HC, Ho LH. A new modified intracerebral hemorrhage score for treatment decisions in basal ganglia hemorrhage: a randomized trial. Crit Care Med 2008;36:2151–2156.
16. Chuang YC, Chen YM, Peng SK, Peng SY. Risk stratification for predicting 30-day mortality of intracerebral hemorrhage. Int J Qual Health Care 2009;21:441–447.
17. Hemphill JC III, Farrant M, Neill TA Jr. Prospective validation of the ICH score for 12-month functional outcome. Neurology 2009;73:1088–1094.
18. Hwang BY, Appelboom G, Kellner CP, et al. Clinical grading scales in intracerebral hemorrhage. Neurocrit Care 2010;13:141–151.
19. Garrett JS, Zarghouni M, Layton KF, Graybeal D, Daoud YA. Validation of clinical prediction scores in patients with primary intracerebral hemorrhage. Neurocrit Care 2013;19:329–335.
20. Baird AE. Improving stroke prognosis. Neurology 2009;73:1084–1085.
21. Woo D, Rosand J, Kidwell C, et al. The Ethnic/Racial variations of Intracerebral Hemorrhage (ERICH) study protocol. Stroke 2013;44:e120–e125.
22. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604–607.
23. Brannen AL II, Godfrey LJ, Goetter WE. Prediction of outcome from critical illness: a comparison of clinical judgment with a prediction rule. Arch Intern Med 1989;149:1083–1086.
24. Finley Caulfield A, Gabler L, Lansberg MG, et al. Outcome prediction in mechanically ventilated neurologic patients by junior neurointensivists. Neurology 2010;74:1096–1101.
25. Navi BB, Kamel H, McCulloch CE, et al. Accuracy of neurovascular fellows' prognostication of outcome after subarachnoid hemorrhage. Stroke 2012;43:702–707.
26. Saposnik G, Cote R, Mamdani M, et al. JURaSSiC: accuracy of clinician vs risk score prediction of ischemic stroke outcomes. Neurology 2013;81:448–455.
27. Boyd EA, Lo B, Evans LR, et al. "It's not just what the doctor tells me": factors that influence surrogate decision-makers' perceptions of prognosis. Crit Care Med 2010;38:1270–1275.

# Clinician judgment vs formal scales for predicting intracerebral hemorrhage outcomes

David Y. Hwang, Cameron A. Dell, Mary J. Sparks, et al.

*Neurology* 2016;86;126-133 Published Online before print December 16, 2015
DOI 10.1212/WNL.0000000000002266