Validation of the NeuroImaging Radiological Interpretation System for Acute Traumatic Brain Injury

Bo Zhou, MD, PhD,† Victoria Y. Ding, MS,‡ Ying Li, MD,* Robyn L. Ball, PhD,‡ Bin Jiang, MD,* Guangming Zhu, MD, PhD,* Derek Boothroyd, PhD,‡ Michael Zeineh, MD, PhD,* Alisa Gean, MD,§|| and Max Wintermark, MD, MAS, MBA**

Purpose: The aim of the study was to refine and validate the NeuroImaging Radiological Interpretation System (NIRIS), which was developed to predict management and clinical outcome based on noncontrast head computerized tomography findings in patients suspected of acute traumatic brain injury (TBI).

Methods: We assessed the performance of the NIRIS score in a prospective, single-center cohort of patients suspected of TBI (n = 648) and compared the performance of NIRIS with that of the Marshall and Rotterdam scoring systems. We also revised components of the NIRIS scoring system using decision tree methodologies implemented on pooled data from the retrospective and prospective studies (N = 1190).

Results: The NIRIS performed similarly to the Marshall and Rotterdam scoring systems in predicting mortality and markedly better in terms of predicting more granular elements of disposition and management of TBI patients, such as admission, follow-up imaging, intensive care unit stay, and neurosurgical procedures. The revised NIRIS classification correctly predicted disposition and outcome in 91.2% (331/363) after excluding patients with other major extracranial traumatic injuries or intracranial nontraumatic injuries.

Conclusions: The present study further demonstrates the predictive value of NIRIS in guiding standardized clinical management and decision-making regarding treatment options for TBI patients.

Key Words: TBI, CT, common data elements, outcome, clinical decision support

(*J Comput Assist Tomogr* 2019;43: 690–696)

Traumatic brain injury (TBI) is a leading cause of death and disability within the spectrum of trauma-related injuries worldwide,[1] with an estimated 69 million individuals affected each year.[2] In the United States, the number of TBI-related emergency department (ED) visits increased from 1.6 million in 2007 to 2.5 million in 2013, with nearly 56,000 TBI-related deaths occurring in 2013.[3] Because neuroimaging can detect and characterize the presence and extent of brain injury, it plays an important role in the triage and management of TBI patients, for which the prognostic value of CT characteristics has been well documented.[4] However, the interpretation of neuroimaging in TBI is not standardized, resulting in variations in terms of clinical practice.

Recently, inspired by the successful use of BI-RADS (Breast Imaging-Reporting and Data System), developed by the American College of Radiology for the communication of results and patient care recommendations in breast imaging, Wintermark et al[5] proposed the NeuroImaging Radiological Interpretation System (NIRIS) for patients suspected of TBI. The NIRIS is an outcome-based, rather than experience-driven, system to standardize the interpretation of noncontrast head computerized tomography (CT) and consolidate imaging findings into different categories of ordinal severity to inform specific patient management actions. When compared with the Marshall[6] and Rotterdam[7] scoring systems, NIRIS performed similarly in terms of survival/death prediction but was clearly superior in terms of predicting specific patient care actions, such as the need for follow-up imaging, intensive care unit (ICU), or neurosurgical procedures. However, the NIRIS system has not been independently validated across a broad range of TBI patients.

The goals of this study were to assess the performance of the NIRIS system in a cohort of patients with a broad range of TBI different from the one in which it was developed and to refine it as needed based on pooled results.

METHODS

Study Design

We prospectively identified and recruited all consecutive patients transported to the Stanford Healthcare Emergency Department by ambulance or helicopter, for whom a trauma alert was triggered per established criteria (http://stan.md/2nw7pfD) and who underwent noncontrast head CT scan because of suspicion of TBI, between December 2015 and April 2017. Our study was approved by our institutional review board.

Imaging Review

Noncontrast head CTs were independently reviewed for TBI imaging common data elements[8] as defined by the National Institute of Health (NIH). Two experienced neuroradiologists assessed, in consensus, the presence/absence of closed head injuries including skull fracture, pneumocephalus, hemorrhage, mass effect, and brain parenchymal injuries. The volumes of each type of hematoma or contusion, as well as the extent of midline shift, were quantified as continuous variables, whereas the extent of subarachnoid hemorrhage, intraventricular hemorrhage, brain edema/swelling, cisternal compression, and hydrocephalus were characterized on ordinal scales.

All patients were categorized based on noncontrast head CT findings per the published NIRIS[5] (Online Table 1, http://links.lww.com/RCT/A85), as well as under the Marshall (Online

From the *Department of Radiology, Neuroradiology Section, Stanford University, Stanford, CA; †Department of Neurology, the Second Medical Centre, National Clinical Research Center for Geriatric Diseases, Chinese PLA General Hospital, Beijing, China; ‡Department of Medicine, Quantitative Sciences Unit, Stanford University, Stanford; §Department of Radiology, Neuroradiology Section, University of California; and ||Zuckerberg San Francisco General Hospital and Trauma Center, San Francisco, CA.

Received for publication April 1, 2019; accepted June 13, 2019.

Correspondence to: Max Wintermark, MD, MAS, MBA, Neuroradiology Division, Department of Radiology, Stanford University, 300 Pasteur Dr, Room S047, Stanford, CA 94305-5105 (e-mail: Max.Wintermark@gmail.com).

This study was in part supported by grant #UL1 TR001085 from the National Institute of Health/National Center for Research Resources.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.jcat.org).

The authors declare no conflict of interest.

DOI: 10.1097/RCT.0000000000000913

Table 2, http://links.lww.com/RCT/A85) and Rotterdam (Online Table 3, http://links.lww.com/RCT/A85) scoring systems.

Clinical Data Collection

Demographic and clinical variables of all patients were extracted from our institution's electronic health record system, including age, sex, mechanism of injury, Glasgow Comas Score, and status at hospital discharge (dead or alive). These general characteristics are summarized in Table 1. We also obtained records of other major extracranial traumatic injuries and other major intracranial nontraumatic injuries (eg, a stroke that would have caused the patient to get involved into a trauma), as well as any complications or follow-up brain imaging during the patient's hospital stay. Imaging characteristics interpreted from noncontrast head CTs were recorded in REDCap,[9] a secure, web-based application designed to support data capture for research studies.

Statistical Analyses

Patient Characteristics

We first compared patients in the original, retrospective NIRIS study to patients in the prospective study with respect to demographics, clinical, and injury characteristics. The standardized difference, also known as Cohen *d*,[10] is a measure of the difference between the distribution of the characteristics in the 2 groups expressed in units of standard deviations. A larger *d* corresponds to a larger difference between the groups, and the magnitude may be interpreted using Cohen's guidelines (0.2 = small difference, 0.5 = medium difference, 0.8 = large difference).[11]

Outcome Definition

Next, we categorized patients from the prospective study into 5 mutually exclusive levels according to their disposition from the ED (Online Table 1, http://links.lww.com/RCT/A85). These ED disposition levels include discharge, admission, advanced care unit stay, neurosurgery, and death. We then summarized clinical and imaging characteristics by ED disposition levels; categorical characteristics were presented as counts and percentages, and continuous characteristics were presented as means with standard deviation (SD) or medians with interquartile range.

Refining the NIRIS Classification Using Conditional Inference Trees

To refine the NIRIS classification, we pooled data collected on patients from the original, retrospective NIRIS study[5] and the prospective study. We used conditional inference trees, implemented using the "party"[12] package in R, to identify which imaging characteristics among NIH TBI imaging common data elements were most strongly associated with each outcome and to identify cutoff values for continuous characteristics. Conditional inference trees have the additional advantages over traditional decision trees of avoiding overfitting and eliminating selection bias toward covariates with many possible splits. We applied this methodology to the pooled cohort, excluding discharged patients, with all other patient outcomes modeled in multivariate form. Rather than considering overall hemorrhage/contusion volume as in the original NIRIS study, we estimated a separate threshold for the volume of subdural hemorrhage because it tends to be larger because it typically occurs in older patients with more pronounced volume loss. All other imaging characteristics were similarly defined as in the original study.

TABLE 1. Demographic, Clinical, and Injury Severity Characteristics in This Prospective Study

Demographics	Overall	Discharged	Follow-up Brain Imaging/Admission	Advanced Care Unit Stay	Neurosurgery	Death
No. patients	**648**	**295**	**190**	**137**	**11**	**15**
Age, mean ± SD, y	50.1 ± 22.1	43.9 ± 19.1	51.5 ± 22.8	58.4 ± 22.5	56.7 ± 27.9	72.2 ± 18.7
Female sex, n (%)	259 (40.0)	119 (40.3)	76 (40.0)	49 (35.8)	6 (54.5)	9 (60.0)
Glasgow Coma Scale, median (Q1–Q3)	15 (15–15)	15 (15–15)	15 (14–15)	15 (14–15)	9.0 (5.5–12.5)	10 (4.5–13)
Mechanism of injury, n (%)						
Fall from height	243 (37.5)	81 (27.5)	78 (41.1)	67 (48.9)	6 (54.5)	11 (73.3)
Motor vehicular collision	340 (52.5)	166 (56.3)	100 (52.6)	65 (47.4)	5 (45.5)	4 (26.7)
Violence	56 (8.6)	43 (14.6)	10 (5.3)	3 (2.2)	0 (0.0)	0 (0)
Other/unknown	9 (1.4)	5 (1.7)	2 (1.1)	2 (1.5)	0 (0.0)	0 (0)
Other major intracranial nontraumatic injury, n (%)	48 (7.4)	4 (1.4)	34 (17.9)	10 (7.3)	0 (0.0)	0 (0)
Other major extracranial traumatic injury, n (%)	245 (37.8)	22 (7.5)	130 (68.4)	79 (57.7)	4 (36.4)	10 (66.7)
Secondary intracranial complications, n (%)	16 (2.5)	0 (0.0)	0 (0.0)	10 (7.3)	3 (27.3)	3 (20.0)
Secondary extracranial Complications, n (%)	36 (5.6)	0 (0.0)	0 (0.0)	22 (16.1)	3 (27.3)	11 (73.3)
Follow-up brain imaging studies, n (%)						
Any noncontrast head CTs	152 (23.5)	0 (0.0)	38 (20.0)	97 (70.8)	11 (100.0)	6 (40.0)
Any contrast-enhanced head CTs	16 (2.5)	0 (0.0)	2 (1.1)	10 (7.3)	3 (27.3)	1 (6.7)
Any intracranial and/or cervical CTAs	48 (7.4)	0 (0.0)	7 (3.7)	34 (24.8)	3 (27.3)	4 (26.7)
Any brain MRIs	34 (5.2)	0 (0.0)	11 (5.8)	16 (11.7)	4 (36.4)	3 (20.0)
Any intracranial and/or cervical MRAs	14 (2.2)	0 (0.0)	3 (1.6)	8 (5.8)	1 (9.1)	2 (13.3)

Percentages are calculated out of 648.

A revised NIRIS classification was thus established and compared with the originally published NIRIS classification, as well as with the Marshall and Rotterdam scoring systems, in the full prospective cohort and in a subcohort without other major intracranial nontraumatic injuries (OMII) or other major extracranial traumatic injuries (OMEI), because these are the main factors besides TBI that can affect the outcomes of trauma patients. Finally, we described all-cause and TBI-related mortality rates in the prospective cohort according to the 3 scoring systems.

All analyses were conducted in the R statistical computing framework,[13] Version 3.3.

RESULTS

Study Population

Six hundred forty-eight adult patients (age: 50.1 ± 22.1 years; 40.0% female) were prospectively enrolled in this study between December 2015 and April 2017, resulting in a pooled cohort of 1190 adult patients (Online Table 4, http://links.lww.com/RCT/A85).

Clinical Data

The clinical characteristics of our study patients are reported in Table 1. Approximately half of the patients were admitted to the ED because of motor vehicle collisions. Nearly 40% of these patients had other major extracranial traumatic injuries, and 7.4% had other major intracranial nontraumatic injuries. Six patients (0.9%) had an external ventricular drain placed, and 6 patients (0.9%) underwent a craniotomy or craniectomy.

Patients in the original NIRIS study[5] and in this prospective study had similar characteristics in many respects (Online Table 4, http://links.lww.com/RCT/A85). Standardized differences in age, sex, OMEI, and follow-up brain imaging studies were low ($d < 0.2$). Disposition at ED between these 2 cohorts were also reasonably similar ($d = 0.25$), although more patients were directly discharged from the ED in this prospective study than in the original, retrospective NIRIS study (45.5% vs 33.6%). There was a moderate difference in the presentation of OMII ($d = 0.33$) and more patients involved in a fall ($d = 1.17$) in this prospective study (Online Table 4, http://links.lww.com/RCT/A85).

Imaging Results

The distribution of the NIH TBI imaging common data elements in this study population stratified by patient outcome is reported in Table 2. The imaging characteristics in the original, retrospective NIRIS study[5] and in this prospective study were similar; both cohorts had 6% of patients with skull fracture, 2% with pneumocephalus, and less than 5% with mass effect.

The performance of the original NIRIS classification[5] in terms of predicting disposition and outcome in this prospective study is demonstrated in Table 3. The original NIRIS classification correctly predicted disposition and outcome in 57.4% (372/648) of all patients and in 87.3% (317/363) of patients without OMEI and OMII. For example, the 295 who were discharged from the ED were correctly classified as NIRIS 0 by the original NIRIS classification. One patient had an old brain injury that was misinterpreted at the time of the ED visit and that prompted a hospitalization; this old brain injury was subsequently recognized as such and received appropriately a NIRIS 0 score. Of note, there were 203 patients with NIRIS 0 and 12 patients with NIRIS 1 who experienced more severe outcomes than expected outcomes. These patients, while NIRIS 0 or 1 for their brain imaging findings, had other complicating factors influencing their care, for example, the OMEI and OMII, as demonstrated by their disappearance when OMEI and OMII were excluded (see second part of Table 3). At the other end of the spectrum, 40 NIRIS 3 or 4 patients experienced better outcomes than those predicted by the original NIRIS classification.

Comparison of Original and Revised NIRIS

The main differences between the original and the revised NIRIS classification (Table 4) include different thresholds for different types of intracranial hemorrhages, the assignment of different severities of hydrocephalus to different NIRIS categories, the addition of pneumocephalus to NIRIS 1 and midline shift >10 mm to NIRIS 4, and the addition of diffuse axonal injury to NIRIS 2.

The performance of the revised NIRIS classification in terms of predicting disposition and outcome in the prospective cohort (n = 648) is demonstrated in Table 5. The revised NIRIS classification correctly predicted disposition and outcome in 60.5% (392/648). After excluding the patients with OMEI and OMII, a correct prediction was observed in 91.2% (331/363) (see second part of Table 5). The revised NIRIS performed better than the original NIRIS classification in predicting disposition both in patients with (60.5% vs 57.4%) and without OMEI and OMII (91.2% vs 87.3%). For example, the 269 patients who were discharged from the ED were still correctly classified as NIRIS 0 by the revised NIRIS classification after OMEI and OMII were excluded. At the other end of the spectrum, only 6 NIRIS 3 patients and 6 NIRIS 4 patients experienced slightly different outcomes than those predicted by the original NIRIS classification, which is the main reason for the improvement of the NIRIS system.

Comparison of Revised NIRIS, Marshall, and Rotterdam Scores for Prediction of All 5 Outcome Categories

The outcomes for each score level for the new NIRIS scoring systems are reported in Table 5. The outcomes for each score level for the Marshall and Rotterdam scoring systems are reported in Online Tables 5 and 6 (http://links.lww.com/RCT/A85).

Marshall scores of 1 and 2 were associated with discharge from the ED, admission, and prolonged hospital stay but could not differentiate among these outcomes. Marshall scores of 5 and 6 were associated with neurosurgical procedures and death (Online Table 5, http://links.lww.com/RCT/A85).

Rotterdam scores were mostly 3 and 4 but did not allow discrimination among the 5 levels of outcome (Online Table 6, http://links.lww.com/RCT/A85).

Comparison of Revised NIRIS, Marshall, and Rotterdam Scores for Prediction of Survival/Death

The proportion of deceased patients generally increased with increasing scores in all scoring systems (Table 6). Nine of 15 patients (2 with NIRIS 3, 7 with NIRIS 4) died from their TBI, whereas 5 patients with NIRIS 0 and 1 patient with NIRIS 2 died from non-TBI–related causes. Sensitivity of revised NIRIS for mortality prediction among patients who died from TBI was 77.8% (7 NIRIS 4 patients of 9 patients who passed away from their TBI) (Online Table 7, http://links.lww.com/RCT/A85). The remaining 2 deceased patients were DNR (do not resuscitate). One of the patients had a subdural hemorrhage, but the volume did not reach 200 mL and midline shift did not reach 10 mm to advance to NIRIS 4. The other patient had extensive intraventricular hemorrhage and mild subarachnoid hemorrhage. These 2 patients might have been saved via neurosurgical intervention if they had not been DNR.

TABLE 2. The NIH TBI Imaging Common Data Elements Observed in This Prospective Cohort

	Overall	Discharged	Follow-up Brain Imaging/Admission	Advanced Care Unit Stay	Neurosurgery	Death
No. patients	**648**	**295**	**190**	**137**	**11**	**15**
Skull fracture, n (%)	39 (6.0)	0 (0.0)	8 (4.2)	24 (17.5)	4 (36.4)	3 (20.0)
Calvarial fracture	26 (4.0)	0 (0.0)	4 (2.1)	17 (12.4)	3 (27.3)	2 (13.3)
Skull base fracture	21 (3.2)	0 (0.0)	4 (2.1)	14 (10.2)	2 (18.2)	1 (6.7)
Pneumocephalus, n (%)	12 (1.9)	0 (0.0)	2 (1.1)	4 (2.9)	4 (36.4)	2 (13.3)
Hemorrhage, n (%)	135 (20.8)	0 (0.0)	28 (14.7)	86 (62.8)	11 (100.0)	10 (66.7)
Epidural hematoma	8 (1.2)	0 (0.0)	2 (1.1)	4 (2.9)	1 (9.1)	1 (6.7)
Volume, mean ± SD, mL	1.7 ± 2.6	0 ± 0	0.6 ± 0.6	1.2 ± 1.4	7.6 ± 0	0.2 ± 0
Subdural hematoma	78 (12.0)	0 (0.0)	7 (3.7)	53 (38.7)	9 (81.8)	9 (60.0)
Volume, mean ± SD, mL	22.0 ± 45.1	0 ± 0	0.5 ± 0.7	11.3 ± 17.0	90.7 ± 89.1	33.2 ± 55.8
Subarachnoid hemorrhage	77 (11.9)	0 (0.0)	18 (9.5)	48 (35.0)	4 (36.4)	7 (46.7)
Intraventricular hemorrhage	9 (1.4)	0 (0.0)	0 (0.0)	7 (5.1)	0 (0.0)	2 (13.3)
Parenchymal hematoma (including hemorrhagic contusions)	18 (2.8)	0 (0.0)	5 (2.6)	9 (6.6)	0 (0.0)	4 (26.7)
Volume, mean ± SD, mL	26.8 ± 50.2	0 ± 0	0.8 ± 1.6	5.7 ± 6.5	0 ± 0	107 ± 56.4
Diffuse axonal injury	7 (1.1)	0 (0.0)	1 (0.5)	5 (3.6)	1 (9.1)	0 (0)
Mass effect, n (%)	28 (4.3)	0 (0.0)	1 (0.5)	10 (7.3)	9 (81.8)	8 (53.3)
Brain edema/swelling	6 (0.9)	0 (0.0)	1 (0.5)	2 (1.5)	2 (18.2)	1 (6.7)
Midline shift	22 (3.4)	0 (0.0)	0 (0.0)	7 (5.1)	8 (72.7)	7 (46.7)
Mean ± SD, mm	7.1 ± 4.6	0 ± 0	0 ± 0	4.1 ± 1.1	7.1 ± 3.6	10.1 ± 6.1
Cisternal compression	10 (1.5)	0 (0.0)	0 (0.0)	1 (0.7)	4 (36.4)	5 (33.3)
Brain herniation/Duret hemorrhages	10 (1.5)	0 (0.0)	0 (0.0)	1 (0.7)	3 (27.3)	6 (40.0)
Hydrocephalus	79 (12.2)	0 (0.0)	20 (10.5)	48 (35.0)	4 (36.4)	7 (46.7)
Mild	53 (8.2)	0 (0.0)	16 (8.4)	34 (24.8)	1 (9.1)	2 (13.3)
Moderate	16 (2.5)	0 (0.0)	2 (1.1)	11 (8.0)	1 (9.1)	2 (13.3)
Severe	8 (1.2)	0 (0.0)	0 (0.0)	3 (2.2)	2 (18.2)	3 (20.0)
Brain parenchymal injuries						
Nonhemorrhagic contusions, n (%)	32 (4.9)	0 (0.0)	3 (1.6)	23 (16.8)	5 (45.5)	1 (6.7)
Volume, mean ± SD, mL	6.4 ± 10.6	0 ± 0	0.5 ± 0.4	5.5 ± 8.6	13.6 ± 19.3	9.5 ± 0

TABLE 3. Distribution of Disposition and Outcome in Our Prospective Study Population Based on the Original NIRIS Scoring System

NIRIS	Discharged	Admitted	Prolonged Stay	Neurosurgery	Death	Total
Prospective study population (n = 648)						
0	295	154	44	0	5	498
1	0	16	12	0	0	28
2	0	18	50	1	0	69
3	0	2	15	3	2	22
4	0	0	16	7	8	31
Total	295	190	137	11	15	648
Prospective study population minus patients with OMEI or OMII (n = 363)						
0	269	1	0	0	0	270
1	0	12	0	0	0	12
2	0	17	30	0	0	47
3	0	2	9	1	0	12
4	0	0	11	6	5	22
Total	269	32	50	7	5	363

The gray-shaded cells are those for which the original NIRIS score predicts appropriately the observed disposition and outcome in our prospective study population.

TABLE 4. The Revised NIRIS

Category	Definition	Patient Management Actions
NIRIS 0	No abnormal finding	Discharge from the ED
NIRIS 1	■ Fracture ± ■ Pneumocephalus ■ Epidural hematoma, subdural hematoma, parenchymal hematoma, or parenchymal contusion <0.5 mL ± ■ Subarachnoid hemorrhage	Follow-up brain imaging and/or admit for observation
NIRIS 2	■ Epidural hematoma, subdural hematoma, parenchymal hematoma, or parenchymal contusion >0.5 mL ± ■ Diffuse axonal injury ± ■ Intraventricular hemorrhage ± ■ Mild or moderate hydrocephalus ± ■ Midline shift 0–5 mm	Admit to a more advanced care unit
NIRIS 3	■ Epidural hematoma, parenchymal hematoma, or parenchymal contusion >15 mL ± ■ Subdural hematoma >50 mL ± ■ Midline shift >5 mm ± ■ Focal herniation	Consider neurosurgical procedure (ventricular drain, burr hole, craniotomy/craniectomy, surgical drainage/evacuation of hematoma)
NIRIS 4	■ Epidural hematoma, parenchymal hematoma, or parenchymal contusion >20 mL ± ■ Subdural hematoma >200 mL ± ■ Severe hydrocephalus ± ■ Midline shift >10 mm ± ■ Diffuse herniation ■ Duret hemorrhage	High risk of TBI-related death

In comparison, TBI-related deaths received Marshall scores ranging from 2 to 6 and Rotterdam scores ranging from 3 to 5 (Online Table 7, http://links.lww.com/RCT/A85).

DISCUSSION

Early and accurate stratification of TBI patients according to the specific findings in their initial brain imaging is critical for their subsequent treatment and management in clinical practice. The NIRIS was developed in a retrospective cohort to provide a reliable granular tool to guide TBI patient care.[5] In this study, we not only assessed the performance of NIRIS in a larger prospective cohort but also revised the NIRIS scoring system to improve its predictive performance in the pooled retrospective and prospective cohorts. The revised NIRIS classification correctly predicted disposition and outcome in 60.5% (392/648) of patients. As other factors besides TBI may affect patient outcomes, our results unsurprisingly showed that the performance of NIRIS dramatically improved (91.2%) upon exclusion of patients with other major injuries.

The revised NIRIS classification performed similarly to the Rotterdam scale and the Marshall scoring system in terms of

TABLE 5. Distribution of Disposition and Outcome in Our Prospective Study Population Based on the Revised NIRIS Scoring System

NIRIS	Discharged	Admitted	Prolonged Stay	Neurosurgery	Death	Total
Prospective study population (n = 648)						
0	295	154	44	0	5	498
1	0	16	12	0	0	28
2	0	20	71	3	1	95
3	0	0	6	3	2	11
4	0	0	4	5	7	16
Total	295	190	137	11	15	648
Prospective study population minus patients with OMEI or OMII (n = 363)						
0	269	1	0	0	0	270
1	0	12	0	0	0	12
2	0	19	43	0	0	62
3	0	0	5	3	1	9
4	0	0	2	4	4	10
Total	269	32	50	7	5	363

The gray-shaded cells are those for which the original NIRIS score predicts appropriately the observed disposition and outcome in our prospective study population.

TABLE 6. Mortality Rates (All-Cause and TBI-Related Only) in the Prospective Study Population According to the Revised NIRIS, Marshall, and Rotterdam Scoring Systems

Score	No. Patients	All-Cause Mortality, n (%)	TBI-Related Mortality, n (%)
NIRIS			
0	498	5/498 (1.0)	0/498 (0)
1	28	0/28 (0)	0/28 (0)
2	95	1/95 (1.1)	0/95 (0)
3	11	2/11 (18.2)	2/11 (18.2)
4	16	7/16 (43.8)	7/16 (43.8)
Marshall			
1	547	5/547 (0.9)	0/547 (0)
2	72	2/72 (2.8)	1/72 (1.4)
3	2	2/2 (100)	2/2 (100)
4	0	0/0 (0)	0/0 (0)
5	5	0/5 (0)	0/5 (0)
6	22	6/22 (27.3)	6/22 (27.3)
Rotterdam			
1	3	0/3 (0)	0/3 (0)
2	561	6/561 (1.1)	0/561 (0)
3	72	3/72 (4.2)	3/72 (4.2)
4	9	4/9 (44.4)	4/9 (44.4)
5	2	2/2 (100)	2/2 (100)
6	1	0/1 (0)	0/1 (0)

predicting mortality; however, it performed markedly better in terms of discriminating among more granular dispositions and elements of care, including discharge, admission and follow-up imaging, prolonged stay, advanced care unit stay, and neurosurgical procedures. These results are in agreement with the original study that described NIRIS.[5] In addition, this is by design because the Marshall and Rotterdam scoring systems were developed for mortality prediction, not more granular elements of patients' disposition and management. In contrast, the NIRIS was designed as an outcome-based (not experience driven) tool with the intent of being more granular in terms of predicting TBI patient care based on the initial imaging, which guarantees a clearer communication of the brain imaging and of the treatment selection in the patient management.

The revised NIRIS system performed better than the original version because of several improvements. First, we refined the NIRIS based on a larger pooled data with almost 1200 patients evaluated for TBI. Second, besides adjusting thresholds for intracranial hemorrhages, we also added new and important criteria to the revised NIRIS to increase classification accuracy. Third, we used a statistical method called conditional inference trees to identify the most relevant features and their thresholds without incurring selection bias. Lastly, data quality and completeness also improved in the prospective study as a result of lessons learned from our earlier study. For example, mechanism of injury was unrecorded in nearly a third of patients in the original study because of legal reasons but almost fully complete in the current study.

We acknowledge several limitations to our study, with the main one being that few patients in our cohort underwent neurosurgery or died, thereby reducing our ability to discriminate between these and less severe outcomes. Compared with the original retrospective population in which NIRIS was developed, the prospective study included even more discharged patients (45.5% vs 33.6%). Consequently, the revised NIRIS classification may benefit from further validation efforts, ideally in a multicenter study that includes more severely injured patients than in our present cohort. Moreover, the classification accuracy of the revised NIRIS was higher than that of the original NIRIS in the prospective cohort but not significantly so at the 0.05 level, suggesting room for refinement through considering additional characteristics associated with TBI.

In summary, we successfully validated the NIRIS in a different prospective cohort. The revised NIRIS classification performed much better than the Marshall and Rotterdam scoring systems in disposition prediction and was also superior to the original NIRIS system. The present study further demonstrates the predictive value of NIRIS in guiding clinical management and decision-making regarding treatment options in TBI patients although multicenter cross validation is still warranted in future.

REFERENCES

1. Rubiano A, Carney N, Chesnut R, et al. Global neurotrauma research challenges and opportunities. *Nature*. 2015;527:S193–S197.
2. Dewan M, Rattani A, Gupta S, et al. Estimating the global incidence of traumatic brain injury. *J Neurosurg*. 2018;1–18.
3. Taylor C, Bell J, Breiding M, et al. Traumatic brain injury-related emergency department visits, hospitalizations, and deaths - United States, 2007 and 2013. *MMWR Surveill Summ*. 2017;66:1–16.
4. Lingsma H, Roozenbeek B, Steyerberg E, et al. Early prognosis in traumatic brain injury: from prophecies to predictions. *Lancet Neurol*. 2010;9:543–554.
5. Wintermark M, Li Y, Ding V, et al. Neuroimaging radiological interpretation system for acute traumatic brain injury. *J Neurotrauma*. 2018;35:2665–2672.
6. Marshall LF, Marshall SB, Klauber MR, et al. The diagnosis of head injury requires a classification based on computed axial tomography. *J Neurotrauma*. 1992;9:S287–S292.

7. Maas AI, Hukkelhoven CW, Marshall LF, et al. Prediction of outcome in traumatic brain injury with computed tomographic characteristics: a comparison between the computed tomographic classification and combinations of computed tomographic predictors. *Neurosurgery.* 2005; 57:1173–1182.
8. Haacke E, Duhaime A, Gean A, et al. Common data elements in radiologic imaging of traumatic brain injury. *J Magn Reson Imaging*. 2010;32: 516–543.
9. Harris P, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381.
10. Austin P. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. *Commun Stat Simul Comput*. 2009;38:1228–1234.
11. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Revised ed. New Jersey: Lawrence Erlbaum Associates, Inc; 1977.
12. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. *J Comput Graph Stat*. 2006;15: 651–674.
13. R: A language and environment for statistical computing. (2016). Available at: https://www.R-project.org/. Accessed on August 5, 2019.