Title

Measures of Intracranial Injury Size Do Not Improve Clinical Decision Making for Children With Mild Traumatic Brain Injuries and Intracranial Injuries

Journal

Neurosurgery, 90(6)

ISSN

0148-396X

Authors

Greenberg, Jacob K
Olsen, Margaret A
Johnson, Gabrielle W
et al.

Publication Date

2022-06-01

DOI

10.1227/neu.0000000000001895

Peer reviewed

Measures of Intracranial Injury Size Do Not Improve Clinical Decision Making for Children With Mild Traumatic Brain Injuries and Intracranial Injuries

Jacob K. Greenberg, MD, MSCI[*]
Margaret A. Olsen, PhD, MPH[‡]
Gabrielle W. Johnson, BA[*]
Ranbir Ahluwalia, BA[§]
Madelyn Hill, MPH[‖]
Andrew T. Hale, MD, PhD[§]
Ahmed Belal, MD[¶]
Shawyon Baygani, BS[¶]
Randi E. Foraker, PhD, MA[‡]
Christopher R. Carpenter, MD, MSc[#]
Laurie L. Ackerman, MD[¶]
Corina Noje, MD[**]
Eric M. Jackson, MD[††]
Erin Burns, MD[‡‡]
Christina M. Sayama, MD, MPH[‡‡§§]
Nathan R. Selden, MD, PhD[‡‡§§]
Shobhan Vachhrajani, MD, PhD[‖‖‖]
Chevis N. Shannon, DrPH, MBA, MPH[‖]
Nathan Kuppermann, MD, MPH[¶¶##]
David D. Limbrick, Jr, MD, PhD[*]

(Continued on next page)

Correspondence:
Jacob K. Greenberg, MD, MSCI,
Department of Neurosurgery,
Washington University School of Medicine,
660 S. Euclid Ave, Box 8057,
St. Louis, MO 63110, USA.
Email: jacobgreenberg@wustl.edu

Received, July 20, 2021.
Accepted, December 5, 2021.
Published Online, March 16, 2022.

BACKGROUND: When evaluating children with mild traumatic brain injuries (mTBIs) and intracranial injuries (ICIs), neurosurgeons intuitively consider injury size. However, the extent to which such measures (eg, hematoma size) improve risk prediction compared with the kids intracranial injury decision support tool for traumatic brain injury (KIIDS-TBI) model, which only includes the presence/absence of imaging findings, remains unknown.

OBJECTIVE: To determine the extent to which measures of injury size improve risk prediction for children with mild traumatic brain injuries and ICIs.

METHODS: We included children ≤18 years who presented to 1 of the 5 centers within 24 hours of TBI, had Glasgow Coma Scale scores of 13 to 15, and had ICI on neuroimaging. The data set was split into training (n = 1126) and testing (n = 374) cohorts. We used generalized linear modeling (GLM) and recursive partitioning (RP) to predict the composite of neurosurgery, intubation >24 hours, or death because of TBI. Each model's sensitivity/specificity was compared with the validated KIIDS-TBI model across 3 decision-making risk cutoffs (<1%, <3%, and <5% predicted risk).

RESULTS: The GLM and RP models included similar imaging variables (eg, epidural hematoma size) while the GLM model incorporated additional clinical predictors (eg, Glasgow Coma Scale score). The GLM (76%-90%) and RP (79%-87%) models showed similar specificity across all risk cutoffs, but the GLM model had higher sensitivity (89%-96% for GLM; 89% for RP). By comparison, the KIIDS-TBI model had slightly higher sensitivity (93%-100%) but lower specificity (27%-82%).

CONCLUSION: Although measures of ICI size have clear intuitive value, the tradeoff between higher specificity and lower sensitivity does not support the addition of such information to the KIIDS-TBI model.

KEY WORDS: Minor head trauma, Intracranial hemorrhage, Risk prediction modeling, Clinical decision support tools, Pediatrics, Child

Neurosurgery 90:691–699, 2022 https://doi.org/10.1227/neu.0000000000001895

Mild traumatic brain injury (mTBI) is one of the most common health problems affecting children.[1-3] Although mTBI can have long-term health effects, its acute evaluation is primarily focused on identifying and appropriately managing the presence of radiographic intracranial injury (ICI). Most children with mTBI and ICI remain neurologically stable, but a minority experience neurological decline and require neurosurgical or other advanced critical care interventions.[4-6] Therefore, matching patient risk to an appropriate level of care is essential for ensuring close neurological monitoring for those at increased risk while avoiding the emotional burden and resource use associated with unnecessary intensive care unit admission.[7,8]

Difficulty in risk-stratifying children with mTBI and ICI has contributed to variable treatment practices that primarily reflect individual physician judgment and institutional culture.[4,9] To advance safe, evidence-based practices, several

ABBREVIATIONS: CDS, clinical decision support; **GLM,** generalized linear modeling; **ICIs,** intracranial injuries; **KIIDS-TBI,** kids intracranial injury decision support tool for TBI; **mTBIs,** mild traumatic brain injuries; **NPV,** negative predictive value; **RP,** recursive partitioning.

Supplemental digital content is available for this article at neurosurgery-online.com.

clinical decision support (CDS) tools have been proposed for managing this population.[4,6,10] One of these tools, the kids intracranial injury decision support tool for TBI (KIIDS-TBI), was developed and externally validated in 2 large multicenter cohorts and is summarized in **Supplemental Figure 1** in the **Supplemental Digital Content**, http://links.lww.com/NEU/C813.[11] The KIIDS-TBI tool builds on long-standing evidence demonstrating the importance of integrating clinical and radiological findings to risk-stratify patients with TBI.[12] The clinical implementation of these findings may improve patient safety and reduce resource use.

Nonetheless, the KIIDS-TBI tool does not include quantitative imaging measures reflecting ICI size (eg, hematoma size), potentially limiting its predictive ability. By comparison, some researchers using smaller study populations have suggested risk-stratification algorithms that incorporate such quantitative measures.[13,14] It is currently unknown whether and to what extent such information improves risk prediction. This question holds significance for CDS intended to be used in emergency settings by multidisciplinary providers with differing levels of neuroimaging experience. Specifically, when radiology reports are relayed verbally in emergency settings, they may not provide detailed measurements of traumatic findings and vary based on local radiology capabilities. Therefore, incorporating quantitative measures of ICI size may preclude CDS use by nonneurosurgeons who may not be comfortable performing such detailed imaging measurements.[15]

Given the potential tradeoff between ease of use and improved risk prediction, the primary objective of this study was to use a large multicenter data set to investigate the value added by including quantitative measures of ICI size in the risk-stratification of children with mTBI and ICI. To aid interpretation of the results of logistic regression modeling,[16] the secondary objective was to visually evaluate the relationship between ICI size and patient risk.

METHODS

Study Data Set

The data set for this analysis was taken from the Pediatric TBI Research Consortium. The details regarding data collection have been published previously.[11] In brief, this data set included children who presented to 1 of the 5 hospital emergency departments with blunt head trauma between 2006 and 2019. Data were collected through retrospective medical record review by site investigators at each participating institution, and a shared operations manual was used to standardize variable definitions. De-identified data were stored in a centralized research electronic data capture database and were subject to further quality control.[17]

Inclusion Criteria

We included children 18 years or younger who presented to a participating emergency department within 24 hours of blunt head trauma, had Glasgow Coma Scale (GCS) scores of 13 to 15, and had ICI on computed tomography or MRI. Consistent with previous studies, ICI was defined as intracranial hemorrhage, cerebral edema, midline shift, pneumocephalus, skull fracture depressed by at least the width of the skull, traumatic infarction, herniation, or venous sinus thrombosis.[11,18] Patients were excluded if they had penetrating head trauma, current brain tumors, premorbid cognitive impairment, a ventricular shunt, coagulopathy, or imaging suggesting a subacute or chronic ICI.

Predictor Variables

The full list of potential predictors with the proportion of missing data is summarized in **Supplemental Table 1** in the **Supplemental Digital Content**, http://links.lww.com/NEU/C813. Imaging variables were captured by site investigators by reviewing radiology reports alone (1 of 5 centers) or in conjunction with primary computed tomography/MRI images (4 of 5 centers) from patients' first neuroimaging scans. Extra-axial, subdural, and epidural hematoma sizes were defined based on the maximum perpendicular distance from the skull. Extra-axial hematomas referred to lesions that could not be confidently distinguished as subdural vs epidural. Cerebral contusion size was defined based on the ABC/2 method, using the maximum hemorrhage dimensions in 3 planes.[19] Fracture depression was measured from the inner table of the skull to the inner cortex of the depressed fracture fragment.

Primary Outcome

The primary outcome was the composite of neurosurgical intervention, death because of TBI, or intubation for more than 24 hours because of TBI. We defined neurosurgical intervention as craniotomy for intracranial hematoma evacuation or lobectomy, repair of cerebrospinal fluid leak, craniotomy for elevation of a depressed skull fracture, intracranial pressure monitor or external ventricular drain placement, or decompressive craniectomy. These outcomes have been reported previously.[4]

Statistical Analysis

Missing predictor variables were imputed using the missForest package in R,[20] and the study data set was randomly divided into training (75%) and testing (25%) data sets based on outcome event rates. There were no missing outcome data. We evaluated 2 statistical approaches to model creation: multivariable generalized linear modeling (GLM) and binary recursive partitioning (RP). These methods were selected because of their high clinical interpretability and strong foundation in the CDS literature.[21,22] For the GLM model, we tested all injury size predictors and GCS score, given its clinical importance and prominent role in the KIIDS-TBI tool. For continuous variables, polynomial and spline transformations were evaluated. In addition, given the large number of potential predictors, we used adaptive lasso to limit and select other

(Continued from previous page)

*Department of Neurological Surgery, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA; ‡Department of Medicine, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA; §Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, USA; ‖Division of Neurosurgery, Dayton Children's Hospital, Dayton, Ohio, USA; ¶Department of Neurological Surgery, Indiana University School of Medicine, Indianapolis, Indiana, USA; #Department of Emergency Medicine, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA; **Department of Anesthesiology and Critical Care Medicine, Division of Pediatric Critical Care Medicine, The Charlotte R. Bloomberg Children's Center, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; ††Neurological Surgery, Johns Hopkins School of Medicine, Baltimore, Maryland, USA; ‡‡Department of Pediatrics, Oregon Health and Science University, Portland, Oregon, USA; §§Department of Neurological Surgery, Oregon Health and Science University, Portland, Oregon, USA; ‖‖Department of Pediatrics, Wright State University, Dayton, Ohio, USA; ¶¶Department of Emergency Medicine, University of California Davis, School of Medicine, Sacramento, California, USA; ##Department of Pediatrics, University of California Davis, School of Medicine, Sacramento, California, USA

TABLE 1. Patient Characteristics

Characteristic	Training cohort		Testing cohort	
	No serious neurological event (n = 1044)	Serious neurological event (n = 82)	No serious neurological event (n = 347)	Serious neurological event (n = 27)
Median age, y (IQR)	4 (10.5)	5.5 (6.8)	4 (10.6)	7 (7)
Age ≥2 y	639 (61.2)	64 (78.0)	217 (62.5)	22 (81.5)
Median (range) hours from injury to ED	1.3 (2.5)	2.8 (3)	1.7 (2.6)	2.6 (3.1)
Patient race				
White	875 (83.8)	71 (86.6)	290 (83.6)	25 (92.6)
Black	124 (11.9)	16 8 (9.8)	39 (11.2)	2 (7.4)
Others	45 (14.4)	3 (3.7)	18 (5.2)	0 (0)
Sex				
Male	660 (63.2)	51 (62.2)	235 (67.7)	20 (74.1)
Female	384 (36.8)	31 (37.8)	112 (32.3)	7 (25.9)
Mechanism of injury				
Fall	534 (51.1)	37 (45.1)	174 (50.1)	12 (44.4)
Moving object struck head	70 (6.7)	12 (14.6)	16 (4.6)	3 (11.1)
Motor vehicle collision	124 (11.9)	7 (8.5)	38 (11.0)	2 (7.4)
Motorcycle/all terrain vehicle accident	54 (5.2)	6 (7.3)	20 (5.8)	1 (3.7)
Others	262 (25.1)	20 (24.4)	99 (28.5)	9 (33.3)
Presenting GCS score				
13	42 (4.0)	7 (8.5)	15 (4.3)	1 (3.7)
14	112 (10.7)	15 (18.3)	46 (13.3)	3 (11.1)
15	890 (85.2)	60 (73.2)	286 (82.4)	23 (85.2)
Noncranial significant injury	119 (11.4)	9 (11.0)	50 (14.4)	6 (22.2)
Severe mechanism of injury	373 (35.7)	38 (46.3)	120 (34.6)	12 (44.4)
Concern for nonaccidental trauma	100 (9.6)	7 (8.5)	42 (12.1)	4 (14.8)
History of post-traumatic seizure	73 (7.0)	5 (6.1)	16 (4.6)	3 (11.1)
CT findings				
Subdural hematoma, n (%)	375 (35.9)	6 (7.3)	129 (37.2)	5 (18.5)
Subdural hematoma median size (IQR)	4 (2)	7.5 (3.3)	3.6 (2)	5.1 (4)
Subarachnoid hemorrhage	306 (29.3)	14 (17.1)	99 (28.5)	5 (18.5)
Epidural hematoma	111 (10.6)	39 (47.6)	39 (11.2)	13 (48.1)
Epidural hematoma size, median (IQR)	6 (4.2)	15 (11.5)	5 (2)	15 (8)
Contusion	163 (15.6)	9 (11.0)	46 (13.3)	6 (22.2)
Contusion median size (IQR)	0.18 (0.33)	0.19 (0.46)	0.17 (0.67)	0.37 (0.50)
Skull fracture depressed ≥the width of the skull	54 (5.2)	42 (51.2)	16 (4.6)	14 (51.9)
Amount of fracture depression, median (IQR)	3 (1.4)	6.2 (3)	3 (4)	5 (1.9)
Midline shift	39 (3.7)	27 (32.9)	13 (3.7)	7 (25.9)
Amount of midline shift, median (IQR)	3 (3)	4 (3.5)	3 (2.0)	4 (4)
Extra-axial hematoma	99 (9.5)	8 (9.8)	36 (10.4)	1 (3.7)
Extra-axial hematoma median size (IQR)	3 (3)	4 (2.1)	4 (2)	2 (NA)

CT, computed tomography; ED, emergency department; GCS, Glasgow Coma Scale; IQR, interquartile range.

potentially influential predictors. Adaptive lasso is a form of penalized regression that has been shown to have desirable statistical properties for variable selection compared with some traditional approaches, such as stepwise selection.[23] We also decided a priori to test for an interaction between the amount of midline shift and epidural hematoma size. Variables with $P < .10$ were retained in the final model. RP involves sequentially identifying the most important predictors and optimal cutoffs, continuing to build a model in a tree-like fashion until further splitting does not improve model accuracy. We graphically evaluated the marginal risk for each imaging variable retained in the final GLM model using the ggeffects package in R.[24] In this context, marginal effects refer to model-based predictions for a variable of interest, holding other model variables constant.

For the RP analysis, we assigned a relative cost of 75:1 for failure to identify a patient who experienced the composite outcome and used a minimum terminal node size of 25 to prevent overfitting. Ten-fold cross-validation was used to select the final model.

We evaluated 3 different risk cutoffs for considering a patient "low risk" (<1%, <3%, or <5%). The performance of each model was investigated by measuring the sensitivity, specificity, and positive and negative predictive values of each threshold. Model calibration was assessed graphically using observed vs predicted values. The performance of the GLM and RP models was compared with the validated KIIDS-TBI model. All analyses were conducted in R version 4.0.1 (R Foundation).[25] The adaptive lasso was performed using the glmnet package,[26] RP was performed using the rpart package,[27] and splines were

TABLE 2. Full Results of the Multivariable GLM Model

Variable	Beta coefficient (95% CI)	Odds ratio (95% CI)	*P*-value
Age 2 y or older	0.89 (0.06 to 1.8)	2.4 (1.1 to 6.0)	.04
Epidural hematoma size	0.38 (0.29 to 0.48)	1.5 (1.3 to 1.6)	<.01
Midline shift amount	0.53 (0.24 to 0.79)	1.7 (1.3 to 2.2)	<.01
Extra-axial hematoma size	0.28 (0.10 to 0.46)	1.3 (1.1 to 1.6)	<.01
Amount of fracture depression	0.91 (0.74 to 1.09)	2.5 (2.1 to 3.0)	<.01
Severe mechanism of injury	0.85 (0.12 to 1.6)	2.3 (1.1 to 4.9)	.02
GCS score = 15	−0.85 (−1.7 to 0.07)	0.43 (0.18 to 1.07)	.06
Epidural size[a] midline shift	[a]	[a]	<.01

GCS, Glasgow Coma Scale.
[a]Given the complexity of representing the interaction between 2 continuous variables, this interaction is displayed graphically in Figure 1.
Estimates are based on the model defined in the training data set.

evaluated using the splines package.[28] *P*-values <.05 were considered statistically significant. Waivers of informed consent and institutional review board approval were obtained at each site.

RESULTS

There were 1126 children included in the training data set and 374 in the test data set, including 82 (7.3%) and 27 (7.2%) with the composite outcome, respectively. The proportion with a serious noncranial injury (11% vs 15%), a serious mechanism of injury (37% vs 35%), and a GCS score of 15 (84% vs 83%) was similar in the training and testing cohorts. Demographic characteristics, clinical history, and imaging findings are summarized in Table 1

GLM Model

Outside of the measures of ICI size and patient GCS score, the following variables were identified as potential model predictors in the adaptive lasso: age younger than 2 years; severe mechanism of injury; and size of the patient's second epidural hematoma (if present). The final GLM model is presented in Table 2 and included epidural hematoma size, amount of midline shift, extra-axial hematoma size, amount of fracture depression, GCS score, severe mechanism of injury, and the interaction between epidural size and midline shift. Using a spline transformation for continuous predictors led to a negligible change in overall model fit (data not shown). The marginal risk for each quantitative predictor (ie, the change in predicted risk holding other model variables constant) is shown graphically in Figure 1. When accounting for other model variables, most predictors showed a small change in risk at small lesion sizes with a steep increase in risk at moderate to large sizes.

RP Model

The final RP model is shown in Figure 2, which shows the most influential variables and model-identified cutoff values for distinguishing risk groups. The model identified a skull fracture depressed by ≥3 mm, an epidural hematoma >4 mm, midline shift ≥2 mm, and extra-axial hematoma size ≥2 mm as key predictors. This model led to 5 risk groups ranging from 0.2% to 49% risk, with 72% of patients classified into the lowest risk category.

Model Predictive Performance

The predictive performance of the GLM and RP models is summarized in Tables 3 and 4. In the training data set, the GLM model and RP models showed similar performance, although the RP model had slightly higher sensitivity across the 3 risk thresholds (94%-98% vs 90%-96%). However, there was a notable drop in the sensitivity of the RP model in the testing data set (89% at all risk levels), whereas there was little change in the GLM model sensitivity (89%-96%). Both models had high specificity in both the training and testing data sets (76%-90% for GLM and 78%-86% for RP in the testing data set). The calibration of both models is shown in **Supplemental Figure 2A and 2B** in the **Supplemental Digital Content**, http://links.lww.com/NEU/C813.

The predictive performance of the KIIDS-TBI model—which uses only categorical predictors—is compared with the GLM and RP models in Tables 3 and 4. In both the training and testing data sets, the KIIDS-TBI model had the highest sensitivity at all 3 risk thresholds (range 93%-100% in the testing data set). Although the magnitude of the difference varied substantially by risk cutoff, the GLM model showed higher specificity (with nonoverlapping 95% CI's) compared with the KIIDS-TBI model at each decision-making cutoff. For example, the specificity of the KIIDS-TBI model was substantially lower than the GLM model at a cutoff of <1% predicted risk (27% vs 76%, respectively). By comparison, the specificity of the KIIDS-TBI and GLM models were similar when using a threshold of <5% risk for clinical decision making (82% vs 90%).

DISCUSSION

We developed and internally-validated 2 risk models that used both quantitative measures of injury size and clinical history to predict the risk of neurosurgical intervention, prolonged intubation

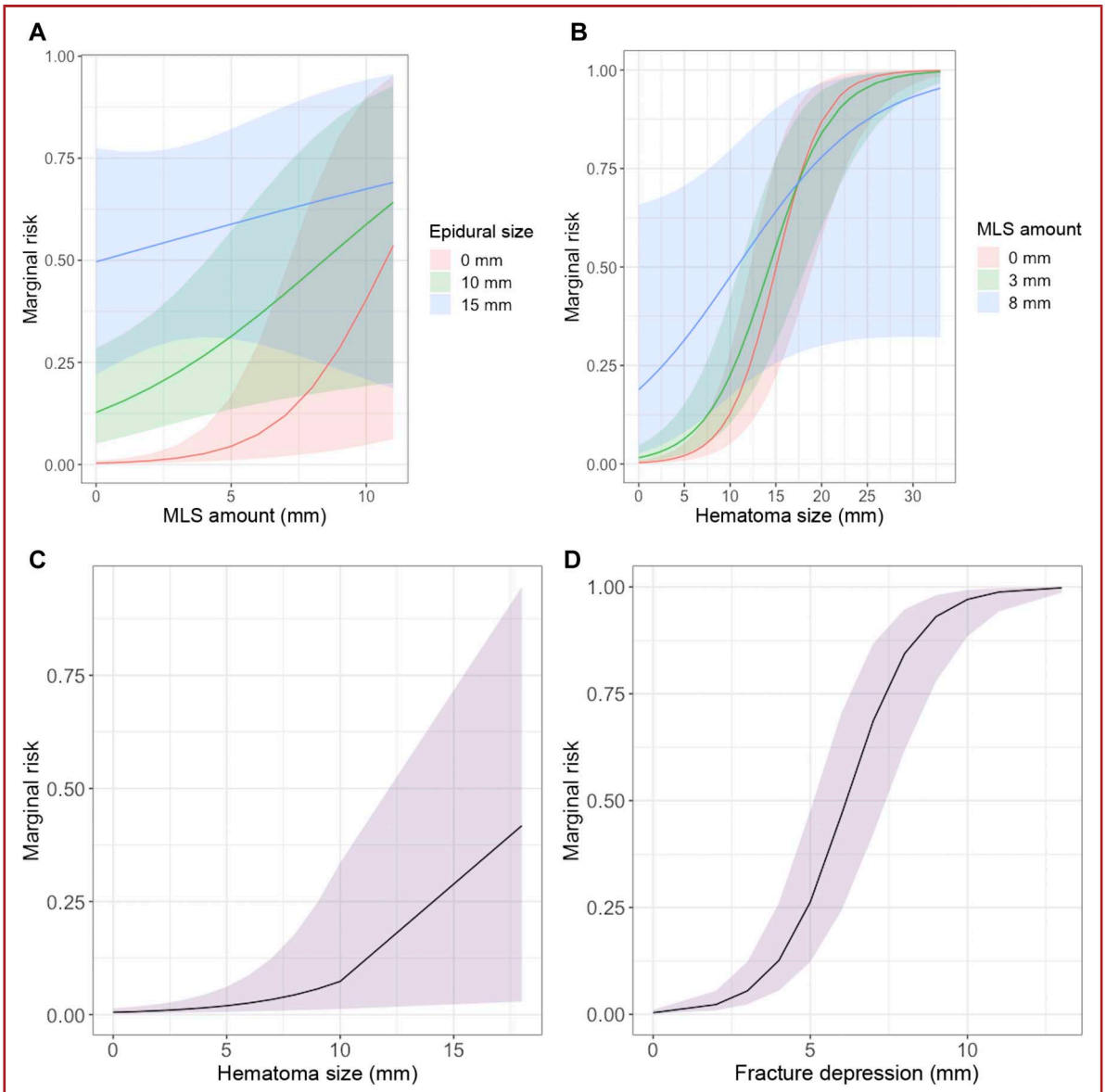

FIGURE 1. *The marginal risk associated with each quantitative imaging predictor included in the final logistic regression model. For each variable, predicted risk is showing across lesion severity while holding other model variables constant.* **A**, *Amount of midline shift as a function of epidural hematoma size;* **B**, *epidural hematoma size as a function of amount of midline shift;* **C**, *extra-axial hematoma size; and* **D**, *amount of skull fracture depression. MLS, midline shift.*

for TBI, or death from TBI in children with mTBI and ICI. We found that the GLM model showed better performance than the RP model in the testing data set. However, compared with the KIIDS-TBI model using only categorical predictors, the GLM model had lower sensitivity but higher specificity, suggesting that quantitative imaging measures added marginal clinical benefit.

A large body of literature has demonstrated that increasing CDS complexity is associated with a lower likelihood of clinical use.[29-31] Despite progress in artificial intelligence, [32,33] automated quantification of intracranial hemorrhage is typically not available in routine clinical workflow. Therefore, clinicians would need to assess quantitative imaging measurements used for CDS. While adding time, this requirement would not affect the tool's accessibility for neurosurgeons. However, as highlighted in recent interviews with multidisciplinary neurotrauma providers, detailed imaging measures may be a barrier for some physicians in specialties less accustomed to reading neuroimaging studies (eg, emergency medicine).[15] Furthermore, decreased use by nonneurosurgeons would be a barrier to using CDS to expand risk knowledge across specialties, 1 potential benefit identified in qualitative analyses.[15]

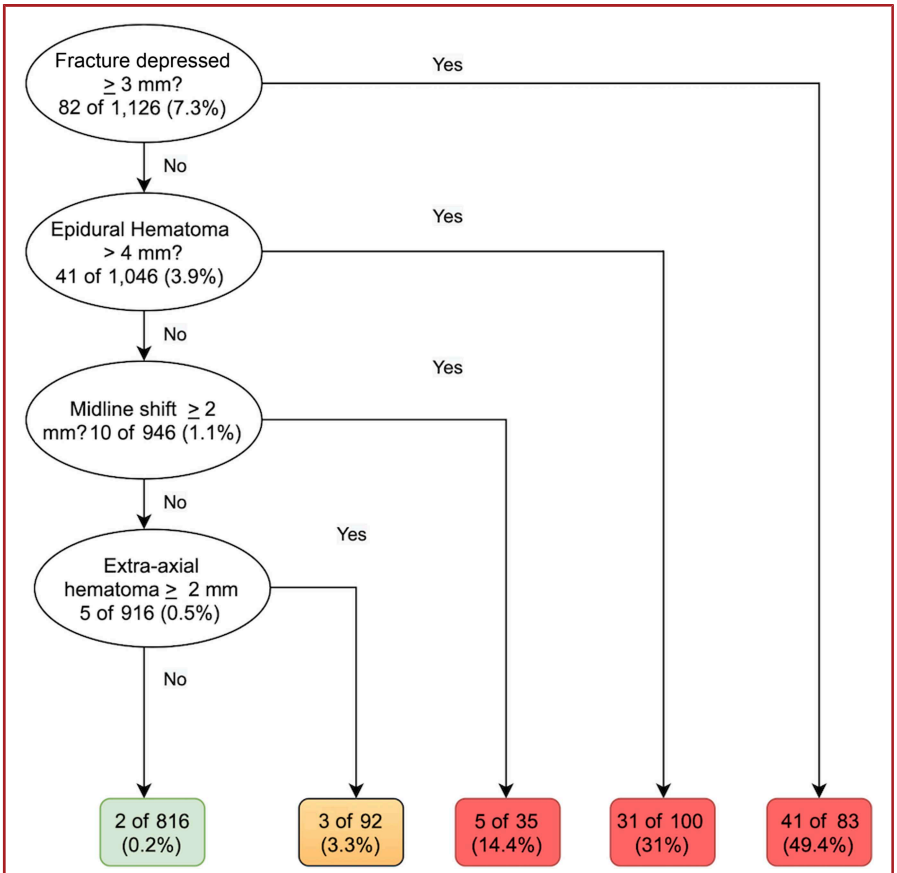

FIGURE 2. *The final recursive partitioning model, which sequentially identified the most influential predictors and cutoff points for each. As the model-building tree moves downward, patients with previously identified risk factors are separated for purposes of identifying subsequent influential predictors. Within each oval is the number of patients who experienced the composite outcome out of the number at risk. Terminal nodes at the bottom of the figure reflect distinct risk groups.*

Consequently, the value added by quantitative imaging measures must be weighed against the challenges created.

Using a large multicenter data set, we found that a GLM model incorporating quantitative measures of ICI size showed higher specificity but lower sensitivity compared with the KIIDS-TBI model, indicating the GLM model may miss more children with the composite outcome. In addition, although the GLM model was validated internally, its performance has not been tested in a new patient population. Taken together, these results suggest that the KIIDS-TBI model is best able to support practical, evidence-based care decisions. In addition, the simplicity of the KIIDS-TBI model may increase its popularity in guiding resource use for multidisciplinary frontline providers in diverse clinical settings.

Although quantitative imaging measures added marginal value to attempts to identify low-risk patients, evaluating the marginal risk predictions shown in Figure 1 can help expand clinical intuition and provide new insights about our results. For example, the effect of midline shift on risk of the composite outcome is substantially lower in patients without any epidural hematoma vs those with a moderate or large (eg, 10 or 15 mm) epidural hematoma. In addition, the wide CI associated with some predictions (eg, extra-axial hematoma larger than 10 mm) is indicative of the small number of patients with those characteristics in the study data set. It is likely that such low numbers at higher lesion volumes also explained why spline transformations for continuous variables did not substantially improve model fit. Therefore, although we used a large data set that included 1500 patients, a much larger population may be needed for precise predictions of infrequent radiographic findings in this population.

Although the relative rarity of some findings limited their value in predictive modeling, our results do not negate the clear clinical importance of considering quantitative imaging measures in appropriate contexts. For example, CDS is not intended to guide the management of a large acute subdural hematoma with mass effect, an undoubtedly a high-risk lesion where clinical experience certainly outweighs evidence-based guidance.

TABLE 3. Performance of the Quantitative GLM Model vs the CHIIDA Model in the Training Data set

	GLM			RP			KIIDS-TBI	
Training data set (n = 1126)	**<1%**	**<3%**	**<5%**	**<1%**	**<3%**	**<5%**	**<1%**	**<3%**	**<5%**
High acuity disposition									
Composite outcome	79	76	75	80	80	77	82	80	78
No composite outcome	235	130	106	230	230	141	752	314	192
Low acuity disposition									
Composite outcome	3	6	7	2	2	5	0	2	4
No composite outcome	809	914	938	814	814	903	292	730	852
Sensitivity (95% CI)	0.96 (0.90-0.99)	0.93 (0.85-0.97)	0.91 (0.83-0.96)	0.98 (0.91-1.0)	0.98 (0.91-1.0)	0.94 (0.86-0.98)	1.0 (0.96-1.0)	0.98 (0.91-1.0)	0.95 (0.88-0.99)
Specificity (95% CI)	0.77 (0.75-0.80)	0.88 (0.85-0.89)	0.90 (0.88-0.92)	0.78 (0.75-0.80)	0.78 (0.75-0.80)	0.86 (0.84-0.89)	0.28 (0.25-0.31)	0.70 (0.67-0.73)	0.82 (0.79-0.84)
PPV (95% CI)	0.25 (0.20-0.30)	0.37 (0.30-0.44)	0.41 (0.34-0.49)	0.26 (0.21-0.31)	0.26 (0.21-0.31)	0.35 (0.29-0.42)	0.10 (0.08-0.12)	0.20 (0.16-0.25)	0.29 (0.24-0.35)
NPV (95% CI)	0.996 (0.99-1.0)	0.99 (0.99-1.0)	0.99 (0.98-1.0)	0.998 (0.99-1.0)	0.998 (0.99-1.0)	0.99 (0.99-1.0)	1.0 (0.99-1.0)	0.997 (0.99-1.0)	0.995 (0.99-1.0)

CHIIDA, Children's Intracranial Injury Decision Aid; GLM, generalized linear modeling; KIIDS-TBI, kids intracranial injury decision support tool for traumatic brain injury; NPV, negative predictive value; PPV, positive predictive value; RP, recursive partitioning.

Limitations

This study has several limitations. First, although each study site attempted to minimize potential bias, the data set relied on retrospective chart review, which may have limited the accuracy with which some variables were recorded.[34] Nonetheless, both model predictors and outcome variables were generally objective imaging findings or major events well-documented in the record, lessening the impact of the retrospective review. In addition, although we evaluated 2 rigorous prediction methods, it is possible that other approaches, such as machine learning algorithms, could have superior performance.[35,36] This possibility should be explored in future work. Third, this analysis was restricted to children with mTBI where neurosurgery is rare, and the results may vary in children with more severe injuries. Finally, although our data set was much larger than those used in previous studies, the sample size limited both the precision of some measurements and also the power of the "testing" validation data set.[37] Future efforts should explore opportunities to expand collaborative research networks in pediatric TBI to address these and other important questions.

CONCLUSION

Analyzing a large, multicenter data set, we found that considering the size of epidural or extra-axial hematoma, midline shift, and amount of fracture depression improved specificity but decreased

TABLE 4. Performance of the Quantitative GLM Model vs the CHIIDA Model in Testing Data set

	GLM			RP			KIIDS-TBI		
Testing data set (n = 374)	**<1**	**<3**	**<5**	**<1**	**<3**	**<5**	**<1**	**<3**	**<5**
High acuity disposition (%)									
Composite outcome	26	25	24	24	24	24	27	26	25
No composite outcome	85	43	36	73	73	46	255	114	62
Low acuity disposition									
Composite outcome	1	2	3	3	3	3	0	1	2
No composite outcome	262	304	311	274	274	301	92	233	285
Sensitivity (95% CI)	0.96 (0.81-1.0)	0.93 (0.76-0.99)	0.89 (0.71-0.98)	0.89 (0.71-0.98)	0.89 (0.71-0.98)	0.89 (0.71-0.98)	1.0 (0.87-1.0)	0.96 (0.81-1.0)	0.93 (0.76-0.99)
Specificity (95% CI)	0.76 (0.71-0.80)	0.88 (0.84-0.91)	0.90 (0.86-0.93)	0.79 (0.74-0.83)	0.79 (0.74-0.83)	0.87 (0.83-0.90)	0.27 (0.22-0.31)	0.67 (0.62-0.72)	0.82 (0.78-0.86)
PPV (95% CI)	0.23 (0.16-0.32)	0.37 (0.25-0.49)	0.40 (0.28-0.53)	0.25 (0.17-0.35)	0.25 (0.17-0.35)	0.34 (0.23-0.47)	0.10 (0.06-0.14)	0.19 (0.13-0.26)	0.29 (0.20-0.39)
NPV (95% CI)	0.996 (0.98-1.0)	0.99 (0.98-1.0)	0.99 (0.97-1.0)	0.99 (0.97-1.0)	0.99 (0.97-1.0)	0.99 (0.97-1.0)	1.0 (0.96-1.0)	0.996 (0.98-1.0)	0.99 (0.98-1.0)

CHIIDA, Children's Intracranial Injury Decision Aid; GLM, generalized linear modeling; KIIDS-TBI, kids intracranial injury decision support tool for traumatic brain injury; NPV, negative predictive value; PPV, positive predictive value; RP, recursive partitioning.

sensitivity compared with the externally validated KIIDS-TBI model. Given this tradeoff and the barriers to use created by quantitative measurements, these results do not support the addition of measures of ICI size to the KIIDS-TBI CDS tool.

Funding

This study was funded by grants from the Agency for Healthcare Research and Quality (1F32HS027075-01A1) and the Thrasher Research Fund (No. 15024). The sponsors had no role in the study design, data collection, data interpretation, manuscript preparation, or the decision to submit the manuscript for publication.

Disclosures

The authors have no personal, financial, or institutional interest in any of the drugs, materials, or devices described in this article. Dr Olsen reports grant funding and consulting fees from Pfizer. Dr Limbrick has received research funds and/or research equipment for unrelated projects from Medtronic, Inc, Karl Storz, Inc, and Microbot Medical, Inc, and has received philanthropic equipment contributions for humanitarian relief work from Karl Storz, Inc and Aesculap, Inc No other authors disclose any competing interests.

REFERENCES

1. Mannix R, O'Brien MJ, Meehan WP III. The epidemiology of outpatient visits for minor head injury: 2005 to 2009. *Neurosurgery*. 2013;73(1):129-134; discussion 34.
2. National Center for Injury Prevention and Control. Report to Congress on Mild Traumatic Brain Injury in the United States: Steps to Prevent a Serious Public Health Problem; 2003.
3. Lumba-Brown A, Yeates KO, Sarmiento K, et al. Centers for Disease Control and Prevention guideline on the diagnosis and management of mild traumatic brain injury among children. *JAMA Pediatr.* 2018;172(11):e182853.
4. Greenberg JK, Yan Y, Carpenter CR, et al. Development and internal validation of a clinical risk score for treating children with mild head trauma and intracranial injury. *JAMA Pediatr.* 2017;171(4):342-349.
5. Greenberg JK, Stoev IT, Park TS, et al. Management of children with mild traumatic brain injury and intracranial hemorrhage. *J Trauma Acute Care Surg.* 2014;76(4):1089-1095.
6. Burns EC, Burns B, Newgard CD, et al. Pediatric minor traumatic brain injury with intracranial hemorrhage: identifying low-risk patients who may not benefit from ICU admission. *Pediatr Emerg Care*. 2019;35(3):161-169.
7. Wang HE, Yealy DM. Distribution of specialized care centers in the United States. *Ann Emerg Med.* 2012;60(5):632.e7-637.e7.
8. Traube C, Silver G, Gerber LM, et al. Delirium and mortality in critically ill children: epidemiology and outcomes of pediatric delirium. *Crit Care Med.* 2017;45(5):891-898.
9. Greenberg JK, Jeffe DB, Carpenter CR, et al. North American survey on the post-neuroimaging management of children with mild head injuries. *J Neurosurg Pediatr.* 2018;23(2):227-235.
10. Ament JD, Greenan KN, Tertulien P, Galante JM, Nishijima DK, Zwienenberg M. Medical necessity of routine admission of children with mild traumatic brain injury to the intensive care unit. *J Neurosurg Pediatr*. 2017;19(6):668-674.
11. Greenberg JK, Ahluwalia R, Hill M, et al. Development and external validation of the KIIDS-TBI tool for managing children with mild traumatic brain injury and intracranial injuries. *Acad Emerg Med.* 2021;28(12):1409-1420.
12. Narayan RK, Greenberg RP, Miller JD, et al. Improved confidence of outcome prediction in severe head injury. A comparative analysis of the clinical examination, multimodality evoked potentials, CT scanning, and intracranial pressure. *J Neurosurg.* 1981;54(6):751-762.
13. Flaherty BF, Moore HE, Riva-Cambrin J, Bratton SL. Pediatric patients with traumatic epidural hematoma at low risk for deterioration and need for surgical treatment. *J Pediatr Surg.* 2017;52(2):334-339.
14. Call L, Qiu Q, Morris J, et al. Characteristics of pediatric patients with traumatic epidural hematomas who can be safely observed: a clinical validation study. *Br J Radiol.* 2020;93(1114):20190968.
15. Greenberg JK, Otun A, Nasraddin A, et al. Electronic clinical decision support for children with minor head trauma and intracranial injuries: a sociotechnical analysis. *BMC Med Inform Decis Mak*. 2021;21(1):161.
16. Van Belle V, Van Calster B. Visualizing risk prediction models. *PLoS One*. 2015;10(7):e0132614.
17. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381.
18. Kuppermann N, Holmes JF, Dayan PS, et al. Identification of children at very low risk of clinically-important brain injuries after head trauma: a prospective cohort study. *Lancet*. 2009;374(9696):1160-1170.
19. Kothari RU, Brott T, Broderick JP, et al. The ABCs of measuring intracerebral hemorrhage volumes. *Stroke*. 1996;27(8):1304-1305.
20. Stekhoven DJ, Bühlmann P. MissForest: nonparametric missing value imputation using random forest. *Bioinformatics*. 2015;28(1):112-118.
21. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med.* 1999;33(4):437-447.
22. Kuppermann N, Dayan PS, Levine DA, et al. A clinical prediction rule to identify febrile infants 60 days and younger at low risk for serious bacterial infections. *JAMA Pediatr.* 2019;173(4):342-351.
23. Zou H. The adaptive lasso and its oracle properties. *J Am Stat Assoc.* 2006;101(476):1418-1429.
24. Lüdecke D. ggeffects: tidy data frames of marginal effects from regression models. *J Open Source Softw.* 2018;3(26):772.
25. *R: A Language and Environment for Statistical Computing*, 4.0.1 ed., R Foundation for Statistical Computing; 2020.
26. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 2010;33(1):1-22.
27. Therneau T, Atkinson B. rpart: recursive partitioning and regression trees; 2019. Accessed February 15, 2022. https://CRAN.R-project.org/package=rpart.
28. Wang W, Yan J. splines2: Regression Spline Functions and Classes; 2020.
29. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA*. 2018;320(21):2199-2200.
30. Bates DW, Kuperman GJ, Wang S, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc.* 2003;10(6):523-530.
31. Ash JS, Sittig DF, Guappone KP, et al. Recommended practices for computerized clinical decision support and knowledge management in community settings: a qualitative study. *BMC Med Inform Decis Mak*. 2012;12(1):6.
32. Lee H, Yune S, Mansouri M, et al. An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. *Nat Biomed Eng.* 2019;3(3):173-182.
33. Kuo W, Häne C, Mukherjee P, Malik J, Yuh EL. Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning. *Proc Natl Acad Sci USA.* 2019;116(45):22737-22745.
34. Worster A, Bledsoe RD, Cleve P, Fernandes CM, Upadhye S, Eva K. Reassessing the methods of medical record review studies in emergency medicine research. *Ann Emerg Med.* 2005;45(4):448-451.
35. Hale AT, Stonko DP, Lim J, Guillamondegui OD, Shannon CN, Patel MB. Using an artificial neural network to predict traumatic brain injury. *J Neurosurg Pediatr.* 2018;23(2):219-226.
36. Filippi M, Cividini C, Agosta F. Deep learning: a turning point in acute neurology. *Lancet Digit Health*. 2020;2(6):e273-e274.
37. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol.* 2005;58(5):475-483.

Supplemental digital content *is available for this article at neurosurgery-online.com.*

Supplemental Table 1. Potential predictors (and % missing) considered during model development.

Supplemental Figure 1. The KIIDS-TBI clinical decision support tool for managing children with mTBIs and intracranial injuries. Risk estimates are based

on the combined derivation and validation cohorts. Reused with permission from John Wiley and Sons, Greenberg et al.

Supplemental Figure 2. Calibration plot comparing observed vs expected risk of the composite outcome across risk groups for the GLM (A) and RP (B) models. The number of patients in each risk group is also shown in the bottom of the figure.

COMMENT

The authors have presented a thorough analysis of the effects of traumatic lesion size on primary neurological outcomes. This study is exploratory to assess if adding more detailed, quantitative imaging findings about lesion size would improve the predictions of the KIIDS-TBI tool, which is a more anatomically descriptive tool. Based on this analysis, injury size alone appears to add little to the value of the tool. The authors appropriately summarize that the more variables to be added into a tool, the more difficult the tool is to use. Because the KIIDS-TBI tool is designed to assist the physician in clinical decision management, it is unlikely that further management refinement (eg, ICU admission, surgery, etc) will result by addition of lesion size to the tool. The study exemplifies the real-world scenario in which the multifactorial aspects of clinical decision making in mild pediatric traumatic brain injury, particularly by frontline, nonneurosurgical clinicians, are likely to be primarily influenced by the presence or absence of neurological and neuroanatomic abnormalities rather than more quantitative, imaged-based measures of neurological injury, the latter being of primary importance to the neurosurgeon in operative decision making in more severe pediatric TBI.

Pedro Aguilar
Martin Weinand
Tucson, Arizona, USA