

Emergency Department Triage of Traumatic Head Injury Using a Brain Electrical Activity Biomarker: A Multisite Prospective Observational Validation Trial

Daniel Hanley, MD, Leslie S. Prichep, PhD, Jeffrey Bazarian, MD, J. Stephen Huff, MD, Rosanne Naunheim, MD, John Garrett, MD, Elizabeth B. Jones, MD, David W. Wright, MD, John O'Neill, MD, Neeraj Badjatia, MD, Dheeraj Gandhi, MD, Kenneth C. Curley, MD, Richard Chiacchierini, PhD, Brian O'Neil, MD, and Dallas C. Hack, MD

ABSTRACT

Objectives: A brain electrical activity biomarker for identifying traumatic brain injury (TBI) in emergency department (ED) patients presenting with high Glasgow Coma Scale (GCS) after sustaining a head injury has shown promise for objective, rapid triage. The main objective of this study was to prospectively evaluate the

From the Brain Injury Outcomes, The Johns Hopkins Medical Institutions (DH), Baltimore, MD; the Department of Psychiatry, New York University School of Medicine (LSP), New York, NY; BrainScope Company, Inc. (LSP), Bethesda, MD; the University of Rochester Medical Center (JB), Rochester, NY; the University of Virginia Health System (JSH), Charlottesville, VA; the Washington University Barnes Jewish Medical Center (RN), St. Louis, MO; the Baylor University Medical Center (JG), Dallas, TX; the University of Texas Memorial Hermann Hospital (EJ), Houston, TX; the Emory University School of Medicine and Grady Memorial Hospital (DW), Atlanta, GA; the Allegheny General Hospital (JO), Pittsburgh, PA; the R. Adams Cowley Shock Trauma Center (NB), Baltimore, MD; the Department of Radiology, University of Maryland (DG), Baltimore, MD; Iatrikos Research and Development Strategies, LLC (KCC), Tampa, FL; the Department of Surgery, Uniformed Services University of the Health Sciences (KCC), Bethesda, MD; R. P. Chiacchierini Consulting, LLC (RC), Gaithersburg, MD; the Detroit Receiving Hospital (BO), Detroit, MI; and Brain Health (DCH), Harpers Ferry, WV.

Received August 29, 2016; revision received January 25, 2017; accepted January 31, 2017.

The authors have no relevant financial information or potential conflicts to disclose.

This study was funded in part by a research contract from U.S. Army, contract W81XWH-14-C-1405, entitled, "Validation of Point-of-Care TBI Detection System for Head Injured Patients." Data acquisition for this study was supported by research grants from BrainScope Co., Inc., to the clinical sites.

JB, JSH, RN, JG, EBJ, DWW, JO, NB, and BO were all principal investigators at clinical data acquisition sites (University of Rochester Medical Center, Rochester, NY; University of Virginia Health System, Charlottesville, VA; Washington University Barnes Jewish Medical Center, St. Louis, MO; Baylor University Medical Center, Dallas, TX; University of Texas Memorial Hermann Hospital, Houston, TX; Emory University School of Medicine and Grady Memorial Hospital, Atlanta, GA; Allegheny General Hospital, Pittsburgh, PA; R. Adams Cowley Shock Trauma Center, Baltimore, MD; Detroit Receiving Hospital, Detroit, MI, respectively), whose institutions received research contracts from BrainScope Company, Inc., to support subject recruitment, consenting and data acquisition. DH led the Brain Injury Outcomes (BIOS) Division of Johns Hopkins, the independent CRO for the trial. He is a member of the Medical Advisory Board of BrainScope Co., Inc., but receives no financial remuneration for this activity. DG served as an independent CT adjudicator, is associate professor and director of interventional neuroradiology at the University of Maryland, and receives salary from BIOS. LSP is employed by BrainScope as the chief scientific officer and is a professor at NYU School of Medicine. LSP holds potential financial interest through patented technology licensed by BrainScope from NYU School of Medicine. DCH currently serves as a consultant to BrainScope Co., Inc., who at the time the study was conducted was coordinator of the Brain Health/Fitness Research Program at the U.S. Army Medical Research and Materiel Command. KCC currently serves as a consultant to BrainScope Co., Inc., who at the time the study was conducted was the Neurotrauma Research Portfolio Manager for Combat Casualty Care Research Program and Defense Health Program at the U.S. Army Medical Research and Materiel Command. RPC served as the independent consulting statistician to BrainScope Co., Inc., for this study.

The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other documentation. In the conduct of research where humans are the subjects, the investigator(s) adhered to the policies regarding the protection of human subjects as prescribed by Code of Federal Regulations (CFR) Title 45, Volume 1, Part 46; Title 32, Chapter 1, Part 219; and Title 21, Chapter 1, Part 50 (Protection of Human Subjects).

The trial (Validation of TBI Detection System for Head Injured Patients [B-AHEAD III]) was registered on ClinicalTrials.gov (NCT02367300; <https://clinicaltrials.gov/ct2/show/NCT02367300?term=BrainScope&rank=5>; date referenced Jun 17, 2016).

Supervising Editor: Peter D. Panagos, MD.

Address for correspondence and reprints: Leslie S. Prichep, PhD; e-mail: leslie.prichep@nyumc.org.

ACADEMIC EMERGENCY MEDICINE 2017;24:617–627.

efficacy of an automated classification algorithm to determine the likelihood of being computed tomography (CT) positive, in high-functioning TBI patients in the acute state.

Methods: Adult patients admitted to the ED for evaluation within 72 hours of sustaining a closed head injury with GCS 12 to 15 were candidates for study. A total of 720 patients (18–85 years) meeting inclusion/exclusion criteria were enrolled in this observational, prospective validation trial, at 11 U.S. EDs. GCS was 15 in 97%, with the first and third quartiles being 15 (interquartile range = 0) in the study population at the time of the evaluation. Standard clinical evaluations were conducted and 5 to 10 minutes of electroencephalogram (EEG) was acquired from frontal and frontal-temporal scalp locations. Using an a priori derived EEG-based classification algorithm developed on an independent population and applied to this validation population prospectively, the likelihood of each subject being CT+ was determined, and performance metrics were computed relative to adjudicated CT findings.

Results: Sensitivity of the binary classifier (likely CT+ or CT-) was 92.3% (95% confidence interval [CI] = 87.8%–95.5%) for detection of any intracranial injury visible on CT (CT+), with specificity of 51.6% (95% CI = 48.1%–55.1%) and negative predictive value (NPV) of 96.0% (95% CI = 93.2%–97.9%). Using ternary classification (likely CT+, equivocal, likely CT-) demonstrated enhanced sensitivity to traumatic hematomas (≥ 1 mL of blood), 98.6% (95% CI = 92.6%–100.0%), and NPV of 98.2% (95% CI = 95.5%–99.5%).

Conclusion: Using an EEG-based biomarker high accuracy of predicting the likelihood of being CT+ was obtained, with high NPV and sensitivity to any traumatic bleeding and to hematomas. Specificity was significantly higher than standard CT decision rules. The short time to acquire results and the ease of use in the ED environment suggests that EEG-based classifier algorithms have potential to impact triage and clinical management of head-injured patients.

The Centers for Disease Control and Prevention estimates that traumatic brain injury (TBI) accounts for over 2.5 million emergency department (ED) visits annually in the United States.¹ ED visits for TBI have increased by 29.1% over the period from 2006 to 2010, a time when ED visits overall increased only 3.6%.² Not included are another 1.6 to 3.8 million annually who sustain sports related TBI and do not seek emergency medical care³ and TBI in the military. TBI treatment is time-sensitive, and early identification is associated with reduced morbidity and improved outcomes.⁴ Robust, quantitative tools for screening do not currently exist in practice. Nonetheless, the rapid, objective, and accurate identification and triage of head-injured persons with such a tool could significantly contribute to improved care and outcome.

Currently, computed tomography (CT) scan is the pragmatically accepted criterion standard for identifying acute intracranial injuries in the ED, although the vast majority of those with mild TBI (mTBI), estimated to be as high as 90%, are found to be negative for clinically important brain injury.^{5–7} Yet many of these patients eventually experience substantial impairment which is underdiagnosed in the ED setting. Clinical decision rules (e.g., New Orleans Criteria [NOC],⁸ Canadian CT Head Rule [CCHR]^{6,9}) currently focus on CT scanning and have high sensitivity

at the expense of very low specificity,^{10–12} contributing to potential long-term health risks associated with over-scanning.^{10,13} Care pathways that reduce the radiation exposure risks and allow for screening of mTBI patients in nonhospital settings could provide substantial benefit to mTBI patients.

Previous studies of traumatic hematomas in mild to moderate head-injured populations, suggest extremely high sensitivity to CT-positive (CT+) cases with measurable blood, using a classification algorithm based on brain electrical activity.^{14,15} In studies comparing brain electrical activity classifiers for TBI with standard practice decision rules for CT in the ED, an electroencephalogram (EEG) marker was reported to have specificity and negative predictive value (NPV) greatly exceeding those of the standard decision rules, while maintaining equivalent sensitivity.^{16,17}

The present, observational, multisite, prospective, clinical trial evaluated the feasibility of using this technology on a handheld device with a disposable headset, at the point of care. It was hypothesized that the classifier algorithm would detect with high accuracy, the likelihood of the patient having a brain injury visible on CT scan (CT+). Coprimary endpoints to test this hypothesis were sensitivity and specificity of the algorithm. Since binary classification does not take into account information related to the patients distance from the threshold, a three-tier (ternary) classification output was also

studied to test the hypothesis that ternary classification would improve accuracy and clinical utility.

METHODS

Study Design and Setting

Following local institutional review board (IRB) approval, this trial included a prospective convenience sample of adult patients presenting to one of the 11 participating U.S. ED sites (Allegheny General Hospital, Pittsburgh, PA; Baylor University Medical Center, Dallas, TX; Detroit Receiving Hospital, Detroit, MI; Emory University School of Medicine and Grady Memorial Hospital, Atlanta, GA; Hartford Hospital, Hartford, CT; R. Adams Cowley Shock Trauma Center, Baltimore, MD; University of Rochester Medical Center, Rochester, NY; University of Texas Memorial Hermann Hospital, Houston, TX; University of Virginia Health System, Charlottesville, VA; Washington University Barnes Jewish Medical Center, St. Louis, MO; Wayne State University Sinai-Grace Hospital, Detroit, MI) following a closed head injury between February and December 2015. Screening occurred in the ED. Subjects were referred for CT scans by the emergency physician in accordance with standard clinical practice. The trial stopping point was event driven and required a minimum of 138 CT+ events. The ratio of CT+ to CT-negative (CT-) was monitored to assure thorough screening of potential subjects. All study subjects met the inclusion/exclusion criteria described below, provided either signed informed written consent or written consent was obtained by proxy (63 cases). The trial (Validation of TBI Detection System for Head Injured Patients [B-AHEAD III]) was registered on ClinicalTrials.gov (NCT02367300; June 17, 2016).

Study Population

The target population consisted of patients with high neurologic function following closed head injury, while in the acute state. Thus, patients between the ages of 18 and 85 years who presented to an ED within 72 hours of suffering head injury, with Glasgow Coma Scale (GCS) of 12 to 15 (at time closest to evaluation by the Ahead 300 device, BrainScope), were candidates for study. Patients were excluded if scalp/skull abnormalities precluded placement of the electrodes on the forehead or if they had advanced dementia, Parkinson's disease, multiple sclerosis,

known seizure or other central nervous system disorders, history of brain tumors, brain surgery, stroke, or evidence of acute psychosis or current substance dependence. In addition, patients with end-stage renal disorder, those requiring airway management, or those receiving procedural sedation at the time of the evaluation were also excluded.

Study Protocol

Clinical Assessments. Subjects were evaluated in the ED using standard practice clinical procedures for each site and its physicians. Evaluations also included the Standardized Assessment of Concussion scale^{18,19} and the Concussion Symptom Inventory.²⁰ While not meant to provide immediate decisions with regard to severity of injury or need for CT scan, these measures were collected only to characterize the symptoms upon presentation to the ED. The NOC and CCHR were computed centrally for study purposes, but were not provided to the clinician nor required during the management process. The use of decision support tools independently by the clinician was neither mandated nor recorded.

EEG Data Acquisition. Five to 10 minutes of eyes closed resting EEG data was recorded using a BrainScope Ahead 300 device and disposable self-adhesive headset. The headset was used to place electrodes on the standard frontal locations of the expanded International 10/20 system, including FP1, FP2, AFz, F7, and F8. All electrode impedances were below 10 k Ω . Amplifiers had a band pass filter from 0.3 to 250 Hz (3 dB points). Data quality was assessed at time of acquisition by artifact algorithms embedded in the device, used to identify and remove any biologic and nonbiologic contamination (e.g., lateral and horizontal eye movement, external electrical noise) and quality was confirmed using additional offline algorithm.²¹

Analysis

EEG Data Analysis. Development of algorithms applied in the validation trial. The patients evaluated in this validation trial were an independent population distinct from the sample used to develop the classification algorithm. The database used for algorithm development was constructed through multiple studies across several years of development, under consistent protocols. Study sites included 20 EDs and 11 colleges and high schools across the United States, with approval from local IRBs. Subjects

were a convenience sample ($n = 2,407$; 36% female, 64% male). Of these subjects 29.1% were controls and 70.9% were TBI patients. TBI patients included males and females between the ages of 15 and 92, who suffered a closed head injury and with a GCS of 8 or higher. The mean GCS of the cohort was 14.9 (median = 15, standard deviation [SD] ± 0.4 , range = 9–15). The mean age of the cohort was 39.5 (median = 36.2, SD ± 17.6 , range = 15.1–91.7) years. All subjects either provided signed informed written consent or written consent was obtained by proxy. Importantly, while similar in demographics and clinical criteria for inclusion, the validation population reported on herein was totally independent of these algorithm development subjects. The details of the data processing approach and sensitivity to detection obtained using this methodology are discussed in published literature^{14,15,22} and is briefly reviewed below.

Structural injury classifier. The binary discriminant classification algorithm applied in this study was derived using a Least Absolute Shrinkage and Selection Operator methodology, which uses a regularized logistic regression model²³ and consists of a weighted combination of selected linear and nonlinear EEG features and selected clinical features that optimally distinguish patients with traumatic structural brain injury visible on CT scan from normal or concussed patients (CT–). This algorithm was then applied prospectively to the validation trial patient population.

The final inputs to the classifier function included EEG and clinical variables. Quantitative EEG (QEEG) features with the highest weights, contributing the most to the classifier function, included: 1) scale-free measures of the total power spectra across regions (especially involving the frontotemporal regions); 2) features reflecting shifts in the frequency spectrum (especially involving the alpha band); and 3) features reflecting disruption in connectivity between regions (especially phase but also coherence). Clinical signs and symptoms believed to be manifestations of the physiologic changes caused by TBI, or risk factors for TBI, also contributed to the classification algorithm, with highest weights for LOC and age.

Derivation of the ternary classification threshold. The ternary classification output was determined using the same development data set as used for the binary (CT+, CT–) analysis, implementing a second threshold (T2) that, together with the binary threshold (T1), define an equivocal zone as a third classification category.

The binary threshold (T1) was derived from algorithm development receiver operating characteristic (ROC) curves, based on discriminant scores for the structural injury classifier function. The distribution of all discriminant scores for the total development population of subjects was used to establish the mean and SD. A conservative interval was specified based on 0.25 times the SD from this population resulting in a second lower threshold (T2) to allow discriminant scores between T1 and T2 to be classified in an equivocal zone. In this way, the normal variance of discriminant scores in the head-injured and non-head-injured population was used to define an expected variance near T1, thus allowing those scores within the range between T1 and T2 to be considered equivocal. The interval of $0.25 \times \text{SD}$ was chosen to limit the percentage of subjects in the interval to less than 10%. In the algorithm development validation population 7% of the patients were found to lie between T2 and T1. These two threshold cut points were applied in an a priori manner to the test population to discriminate CT+ from CT–.

Analysis of trial data. All EEG data processing was completed offline to maintain data acquisition blind to the clinical presentation and to blind the classification results at the clinical site. It is important to note that since the classification algorithm was finalized a priori, only those specific features used in the algorithm were extracted from the independent validation population as part of the algorithm calculation used to classify each subject's EEG.

CT Adjudication for Clinical Truth

In all cases the determination to receive a CT scan was made by the site ED physician, according to standard of care. The centrally adjudicated results of the routine CT served as the pragmatic reference standard. CT scans, as DICOM images, were deidentified and transferred for adjudication independent of the test EEG data results. A positive finding was prospectively defined as an adjudicated core laboratory reading of a subject's clinical CT with the determination of the presence of intracranial blood. To address the potential differences between neuroradiologic reads of the CT scans across sites, adjudication of clinical truth followed a rigorous and quantitative procedure involving sequential evaluation by imaging specialists and physician specialist readers with image-based initial independent determination of CT+ or CT– and then adjudication of discrepant readings and adjudicated

unanimity for final determinations as the ultimate gold standard for image-based truth. Additionally, blood volume measurement was determined using OsiriX (which included a human segmentation capability).²⁴ A blood volume greater than or equal to 1 mL was used as the threshold to identify a subset of patients with well-formed hematomas that may represent a greater clinical risk of injury. CT+ findings included anatomic description of location such as subarachnoid hemorrhage (SAH), subdural, epidural, and/or intracranial hematomas. Positive findings did not include extracranial injuries (such as scalp lacerations/foreign bodies, soft tissue swelling), facial injuries (fractures, blood in sinuses), and nontraumatic abnormalities (such as sinusitis). If the patient received more than one CT scan during their ED evaluation, the scan closest in time to the Ahead 300 evaluation was adjudicated to determine clinical truth.

Since site IRBs would not allow CT scans in subjects where it was not determined to be clinically necessary, and in keeping with multiple published guidelines for reduction of unnecessary radiation exposure,²⁵ these patients were considered CT- only if the NOC was applicable (that is, they had a GCS = 15 and LOC or posttraumatic amnesia) but were negative for all NOC symptoms (i.e., the NOC was negative) and if study follow-up evaluations (at 72–96 hours postinjury) confirmed no exacerbation of symptoms or return for further neuroimaging or treatment.²² This procedure was applied in the Food and Drug Administration–cleared Ahead 100/200 and Ahead 300 validation trials.

Data Analyses

Planned Analyses. The primary objective of this study was to validate the clinical utility of the BrainScope Ahead 300 device for the acute identification of structural brain injuries visible on CT in the TBI population, following closed head injury. The coprimary endpoints for this study were the sensitivity/specificity pair for identification of CT+ cases by the Ahead 300 classification algorithm relative to adjudicated CT truth. A secondary objective was to evaluate the clinical utility of creating a system of classification into three tiers, likely CT+, equivocal [require close observation], or likely CT-.

Power to Observe. Power analyses for sensitivity determined that with 80% power the number of CT+ patients required was 138 for a one-sided alpha of 0.05. For the specificity at least 543 CT- subjects were needed for a one-sided alpha of 0.05. If either the criterion

standard or the study test where found to be of indeterminate status on blind review of test quality they were eliminated from the analysis; thus the paired data set had no missing data. All data analyses were performed by an independent biostatistician. Data acquisition and compliance to the protocol was independently monitored by a Contract Research Organization (Brain Injury Outcomes [BIOS] Division, Johns Hopkins University). The data were analyzed in accordance with a prespecified statistical report and analysis plan prepared by the independent biostatistician prior to conduction of the study, to yield multiplicity-adjusted test statistics for primary and secondary endpoints. The analyses were done with StatXact Version 8 or later. Creation of analysis data sets was done with SAS version 9.2 or later.

Role of the Funding Sources. The funding source worked with clinical sites to develop the protocol to address the endpoints of the U.S. Army–funded contract. The final protocol was approved by each site’s IRB and the U.S. Army Human Research Protection Office prior to study activation. BrainScope supported data collection by research contracts to the clinical sites and provided training in use of the Ahead device and data acquisition. Funding supported engagement of an independent CRO (BIOS, Johns Hopkins) who sequestered all data throughout the study, keeping BrainScope blinded to the reference data, with the exception of one clinical coordinator who worked closely with the CRO and clinical sites but remained blinded to the Ahead data. The corresponding author, LSP, was blinded to the data until after the independent statistician (RPC) completed all analyses of primary and secondary endpoints and worked with clinical site co-authors in the writing of the report and in interpretation of the data.

RESULTS

Characteristics of the Study Participants

A total of 720 closed head-injured subjects were enrolled. Figure 1 shows the diagram of patients eligible for enrollment and the study population. Symptomatology classification demonstrated groups of subjects with moderate (30.52%), mild (47.94%), and no symptoms (21.54%).

Of the 720 subjects, 156 were CT+ and 564 were CT-. Of the 564 CT- subjects, 155 (27%) did not have a CT ordered under standard practice guidelines and were deemed CT- for purpose of analyses

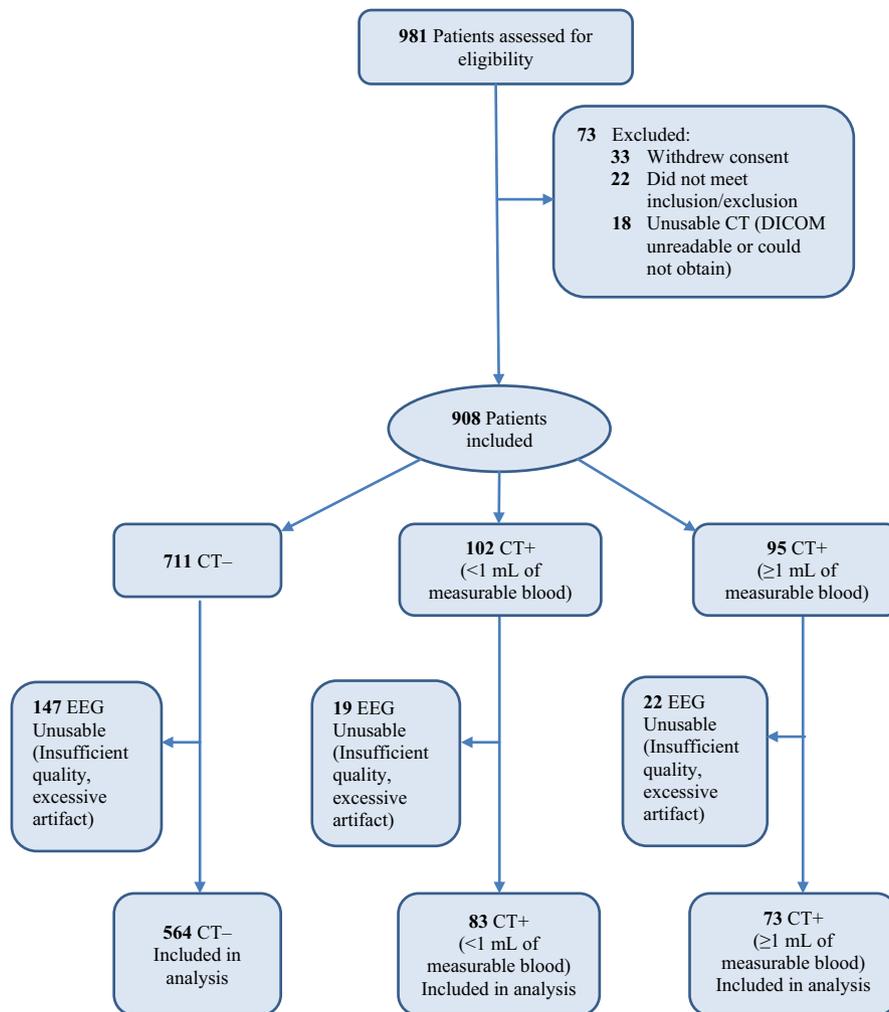


Figure 1. Flow chart showing patients screened for eligibility to the study through to final study population evaluated with details at each node. EEG = electroencephalogram. [Color figure can be viewed at wileyonlinelibrary.com]

according to the procedure described in detail above. Thus, the determination of CT⁻ included a follow-up (phone within 72–96 hours or 30-day medical record review) as verification that a subject released from the ED did not have a structural injury that either evolved or was initially missed. A total of 37.5% of subjects were hospitalized for observation, 8.2% were admitted to an ED observation unit, while 54.3% were released from the ED to return home. One subject required neurosurgical intervention for a depressed temporal bone fracture; this subject was identified independently as CT⁺ by both the core CT laboratory and the test device.

Table 1 presents the characteristics of the CT⁻ and CT⁺ populations of patients. Significant differences were found for age, with the CT⁺ population having a higher mean age, likely reflecting the increased vulnerability/susceptibility to TBI following head injury in the elderly population. GCS also showed a significant difference between groups.

However, with the small range of values, identical medians, a very small SD, and the first and third quartiles being 15 (IQR = 0), this difference is considered to be not clinically meaningful. Table 2 shows the mechanisms of injury, with the largest percentage of both populations coming from motor vehicle accidents and falls.

Other Information. There were one adverse event (burning sensation on forehead) and five serious adverse events (readmission to the hospital) reported during the trial. All serious adverse events were unrelated to the device or the subject's participation in the trial.

Test Results

The contingency table and performance metrics for classification as likely CT⁺ or likely CT⁻ are shown in Table 3. Sensitivity was found to be 92.3% and specificity 51.5%. It was additionally noted that

Table 1
Demographics and Clinical Information for CT- and CT+ patients

Baseline Characteristic	CT-	CT+	t-value	p-value
Age (y)	41.21 (±17.52), 38.04 (18.05 to 84.11), n = 564	52.24 (±19.94), 55.23 (18.00 to 85.62), n = 156	-6.27	<0.0001
GCS (at time of evaluation)	14.98 (±0.17), 15.00 (12 to 15), n = 564	14.91 (±0.37), 15.00 (13 to 15), n = 156	2.33	0.0211
Time between CT scan and BrainScope assessment	6.43 (±11.38) 2.52 (-27.90 to 65.90), n = 409	8.29 (±11.96), 5.58 (-25.40 to 57.38), n = 156	-1.67	0.0951
Sex (% male)	57.3% (323/564)	73.1% (114/156)	—	—

Data are reported as mean (±SD), median (range) for each characteristic and the t- and p-value for each. It is noted that only one subject had GCS less than 13, and that the first and third quartiles equal to 15 (IQR = 0).
CT- = CT negative; CT+ = CT positive; GCS = Glasgow Coma Scale.

Table 2
Mechanism of Injury for the CT- and CT+ Patients

Mechanism of Injury	CT-	CT+
Motor vehicle collision	40 (224/564)	21 (32/156)
Motorcycle/bike accident	8 (48/564)	6 (10/156)
Assault	11 (61/564)	18 (28/156)
Sports related	2 (14/564)	1 (2/156)
Fall related	29 (163/564)	45 (70/156)
Struck by vehicle	5 (27/564)	3 (4/156)
Other	5 (27/564)	6 (10/156)

Data are reported as percentage in each category for each group.
CT- = CT negative; CT+ = CT positive.

specificity was demonstrated to scale with severity of clinical functional impairment with specificities of 76.7, 58.8, and 22.2% for none, mild, and moderate functional impairment, respectively. It is noted that the scoring for severity of functional impairment was performed retrospectively, in a blind manner, using an algorithm that was based on the report of the presence and severity of clinical signs and symptoms.

There were 12 false-negative (FN) classifications resulting in a FN rate of 7.7%. Further, eight of these 12 patients were classified in the equivocal zone in the ternary classification, indicating that these patients were close to the threshold for the binary classification, therefore suggesting need for further observation or evaluation of this group, lowering the false-positive rate (classified as clearly negative) to 2.5%. None of the FNs required surgery or returned to the hospital for exacerbation of symptoms or additional neuroimaging. Detailed study of the clinical/medical characteristics of the FNs revealed that all had GCS = 15, none had any focal neurologic signs, six reported LOC, none of the patients had an epidural hematoma (EDH), and half of the CT+ findings included SAH

Table 3
Contingency Table for Classification Accuracy of Ahead 300 Algorithm for Likelihood of CT+/CT-

AHEAD 300 Classification	Clinical Class		Total
	CT-	CT+	
Likely CT-	291	12	303
Likely CT+	273	144	417
Total	564	156	720

Se = $100 \times 144/156 = 92.31\%$ (87.84%–95.50%)*
Sp = $100 \times 291/564 = 51.60\%$ (48.05%–55.13%)*
NPV = $100 \times 291/303 = 96.04\%$ (93.18%–97.94)†
PPV = $100 \times 144/417 = 34.53\%$ (29.97%–39.31)†

CT- = CT negative; CT+ = CT positive; NPV = negative predictive value; PPV = positive predictive value; Se = sensitivity; Sp = specificity.
*Two-sided 90% CIs (one-sided 95% lower limit is same as the lower two-sided 90% limit).
†Two-sided 95% CIs. NPV and PPV estimates are relevant for the prevalence in study (21.67%).

(alone or with a small hematoma). No relationship to specific locations of the CT findings were found in this FN group. Further, a wide age range and time of evaluation relative time of injury suggested that neither were contributing factors.

Using the prevalence of CT+ in this study (21.7%, 156/720), NPV was found to be 96.0% (93.2%, 97.9) and PPV was found to be 34.5% (30.0%, 39.3%). It is noted that the literature supports a prevalence rate for equivalent, general populations to be approximately 10%;²² thus NPV and PPV was also computed at this prevalence rate where NPV was 98.4% and PPV was 17.5%.

The ROC curve for performance of the structural injury classifier in this independent, validation population had an area under the ROC curve (AUC) of 0.82 and is shown in Figure 2. The circle on the ROC curve indicates the binary classification threshold (T1) determined in prior algorithm development

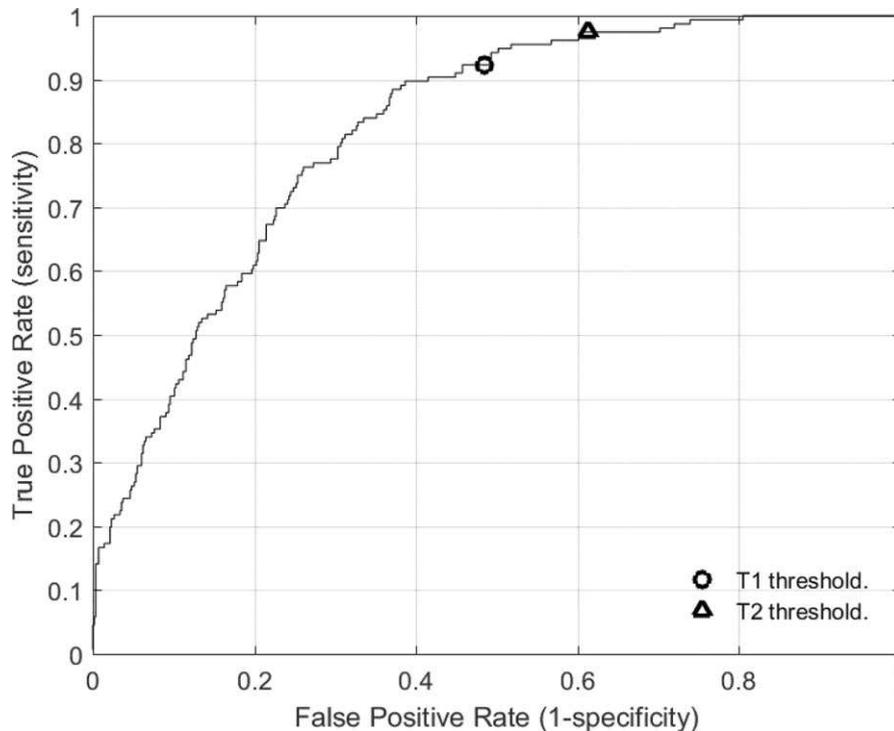


Figure 2. ROC curve for performance of the classifier in the independent validation population. The ROC curve shows the a priori determined threshold (T1, *circle*) for binary classification and the second threshold (T2, *triangle*) used to define the equivocal zone for ternary classification. The AUC for this ROC is 0.82. AUC = area under the ROC curve; ROC = receiver operating characteristic. [Color figure can be viewed at wileyonlinelibrary.com]

studies. Classifications in this independent validation trial were made relative to this a priori determined threshold. It should be noted that the AUC for a classifier using only clinical features as inputs was 0.75. Given that AUC ranges from 0.5 to 1.0, this is a significant difference in AUC, representing a relative increase of 28% in overall classifier performance using the structural injury classifier (EEG plus clinical), supporting the clinical value added of the core EEG features in the classification algorithm.

It is noted that homogeneity across the 11 clinical sites was tested for both sensitivity and specificity. There was little evidence of lack of homogeneity across sites ($p = 0.5718$ for sensitivity and $p = 0.3858$ for specificity), justifying the pooling of data across sites.

Three-tier Classifications

Combining positive classification and equivocal classification, sensitivity becomes 97.4% (152/156, 95% confidence interval [CI] = 93.6%–99.3%); specificity becomes 38.7% (281/564, 95% CI = 34.6%–42.8%); NPV becomes 98.2% (218/222, 95% CI = 95.5%–99.5%); and PPV becomes 30.5% (152/498, 95% CI = 26.5%–34.8%). Note that combining the CT– and the equivocal zone generates the same results as

shown above for the binary classifier (see Table 3), as the second threshold (T2) was below the binary threshold (T1) of the classifier. The second threshold (T2) is shown as a triangle on the ROC curve seen in Figure 2.

Additional Diagnostic Efficacy

Seventy-three of the CT+ patients were found to have measurable blood > 1 mL shown on CT DICOM images (including SAH, EDH, subdural hematoma [SDH], and intracranial hematoma [ICH] traumatic hemorrhages), representing 46.8% of the total CT+ population. The sensitivity of the binary classifier to this subgroup was 93.2% (67/73, 95% CI = 87.8%–95.5%) for the binary classification and 98.6% (72/73, 95% CI = 92.6%–100.0%) for the three-tier classification. The CIs for this endpoint are not adjusted.

DISCUSSION

This multisite validation study demonstrated that a hand-held, noninvasive, easy-to-use device can objectively assess TBI in the ED with 92.3% sensitivity, 51.6% specificity, and 96.0% NPV, in an independent test population of 720 mild-moderate head-injured patients, using a classification algorithm derived from

a large, independently acquired population of head-injured patients seen in the EDs across the country. CT+ cases were independently adjudicated to establish clinical truth and included any structural intracranial brain injury as visible on CT scan (including SAH, SDH, EDH, ICH, and combinations of these). The importance of validating the algorithm in the acute state (within 72 hours of injury), in a population who presents with mild neurologic impairment (97% with GCS = 15), supports the potential impact of the methodology in those patients where triage is often difficult, yet time of clinical course of utmost importance.

Binary classification suffers from the fact that the distance of the subject from the threshold is not taken into consideration. Incorporation of an “equivocal zone” indicating when a classification is near the threshold, that is, higher than normal, but not high enough to be considered positive (e.g., “prehypertension,” “prediabetes”), was added and the clinical utility assessed relative to the binary classification results. Higher sensitivity (97.4%) was obtained for three-tier classification of likely CT+, likely CT-, or equivocal, when equivocal is treated as a positive result. Sensitivity to the subset of traumatic hematomas with ≥ 1 mL of blood was 93.2% (binary) and 98.6% (ternary), demonstrating substantial sensitivity to traumatic brain injuries where risk related to FNs is highest. It is also important to note that the accuracy of identification of hematomas was not related to distance from recording electrodes (data not shown) and was obtained for bleeds at a reliably detectable volume (≥ 1 mL). Since FNs were of highest concern, it is noted that the small number of FNs did not include any patients who required neurosurgery. Study results demonstrate the enhanced sensitivity performance of the three-zone classification, providing additional clinically important information about subjects whose classification is close to the binary threshold, thereby identifying a group of patients who might require further observation.

These high levels of sensitivity were obtained along with overall specificities of 51.6% (with binary classification) and 38.7% (with three zones). Specificity was also found to be inversely related to degree of functional impairment reported. The highest specificity (76.7%) was reported in those with little to no functional impairment. The specificity of NOC applied to this study population was found to be 8.6%, similar to that reported in the literature¹¹ and manyfold lower than the 51.6% specificity obtained in this study, with approximately equivalent sensitivity. Applying the

CCHR,⁹ another standard decision rule for such population, specificity of 31% was obtained, while higher than the NOC, is still significantly below that of the Ahead 300 validation results. It is noted that these reduced specificities were obtained with approximately equivalent sensitivity to that obtained with the Ahead 300 classification algorithm, with 97 and 94%, respectively, for the NOC and the CCHR.

Negative predictive value of 96% (for binary classification) and 98.2% (for ternary classification) reported in this trial can increase confidence in the clinical decision path in patients with minor or no possibly aiding in the reduction of overscanning. Likewise, a positive finding could result in scanning a patient with a very mild presentation but who might have a structural injury.²⁶ This technology is not meant to replace the CT scan in patients with head injury, but provides the clinician with additional information to facilitate routine clinical decision making. Demonstrations of real-world sensitivity in epidemiologic data sets will eventually inform the precise level of sensitivity for CT+, which would best inform utilization in practice.^{27–30}

LIMITATIONS

The study was limited to an adult population; further studies are under way to expand into the pediatric population. The total number of individuals and brain lesions is still small compared to the yearly number of traumatic lesions. Estimates from a well-designed epidemiologic evaluation will give even more robust measures of sensitivity and specificity. Perhaps such estimates will be useful for long-term public health purposes. Additionally, clinical sites did not include urgent care facilities where such capabilities could be clinically important. All analyses in this study were conducted offline. Future studies need to explore mitigation of conditions that may interfere with data quality under “clinical use” versus the “research setting.” Future studies are needed to further evaluate the impact of physicians using such data in real-time acute evaluation of mTBI patients.

CONCLUSIONS

Using an algorithm based on quantitative brain electrical activity has been shown to have potential as a biomarker of traumatic structural brain injury, with high sensitivity and high negative predictive value for acute traumatic bleeding on computed tomography scan. This algorithm

based on advanced signal processing methods and machine learning technology demonstrated enhanced clinical utility of electroencephalogram, well beyond that of conventional methods. The short time to acquire results and the ease of use in the ED suggest utilization of this type of device in other urgent care environments, at “field” or other underresourced areas, as an adjunct to traumatic brain injury assessment where imaging is unavailable. In summary, objective devices with clinically meaningful sensitivity, specificity, and negative predictive value could substantially alter the landscape of mild traumatic brain injury management as an adjunct to the clinical diagnostic pathways, to assist in diagnosis of bleeding and of nonbleeding brain injuries.

The authors acknowledge the work of W. Andrew Mould, Hasan Ali, and Krista Vermillion of BIOS (Brain Injury Outcomes, The Johns Hopkins Medical Institutions) for their dedicated contributions to the study. We also acknowledge the contributions of Dr. Tamara Espinoza and all research staff at all the clinical sites for their efforts toward conducting of this study.

References

- Bell JM, Breiding M, Jenkins EL, et al., editors. Report to Congress on Traumatic Brain Injury in the United States: Epidemiology and Rehabilitation. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control; Division of Unintentional Injury Prevention, 2015.
- Marin JR, Weaver MD, Yealy DM, et al. Trends in visits for traumatic brain injury to emergency departments in the United States. *JAMA* 2014;311:1917–9.
- Langlois JA, Rutland-Brown W, Wald MM. The epidemiology and impact of traumatic brain injury: a brief overview. *J Head Trauma Rehabil* 2006;21:375–8.
- Moppett I. Traumatic brain injury: assessment, resuscitation and early management. *Br J Anaesth* 2007;99:18–31.
- Korley FK, Kelen GD, Jones CM, et al. Emergency department evaluation of traumatic brain injury in the United States, 2009–2010. *J Head Trauma Rehabil* 2016;31:379–87.
- Mower WR, Hoffman JR, Herbert M, et al. Developing a decision instrument to guide computed tomographic imaging of blunt head injury patients. *J Trauma Acute Care Surg* 2005;59:954–9.
- Stiell IG, Wells GA, Vandemheen K, et al. Variation in ED use of computed tomography for patients with minor head injury. *Ann Emerg Med* 1997;30:14–22.
- Haydel MJ, Preston CA, Mills TJ, et al. Indications for computed tomography in patients with minor head injury. *N Engl J Med* 2000;343:100–5.
- Stiell IG, Wells GA, Vandemheen K, et al. The Canadian CT Head Rule for patients with minor head injury. *Lancet* 2001;357:1391–6.
- ACEP Announces Lists of Tests as Part of Choosing Wisely Campaigns. Irving, TX: American College of Emergency Physicians; 2013.
- Papa L, Stiell IG, Clement CM, et al. Performance of the Canadian CT Head Rule and the New Orleans Criteria for predicting any traumatic intracranial injury on computed tomography in a United States Level I trauma center. *Acad Emerg Med* 2012;19:2–10.
- Ro YS, Shin SD, Holmes JF, et al. Comparison of clinical performance of cranial computed tomography rules in patients with minor head injury: a multicenter prospective study. *Acad Emerg Med* 2011;18:597–604.
- Smith-Bindman R, Lipson J, Marcus R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Arch Intern Med* 2009;169:2078–86.
- Hanley DF, Chabot R, Mould WA, et al. Use of brain electrical activity for the identification of hematomas in mild traumatic brain injury. *J Neurotrauma* 2013;30:2051–6.
- Prichep LS, Naunheim R, Bazarian J, et al. Identification of hematomas in mild traumatic brain injury using an index of quantitative brain electrical activity. *J Neurotrauma* 2015;32:17–22.
- Ayaz SI, Thomas C, Kulek A, et al. Comparison of quantitative EEG to current clinical decision rules for head CT use in acute mild traumatic brain injury in the ED. *Am J Emerg Med* 2015;33:493–6.
- O’Neil B, Prichep LS, Naunheim R, et al. Quantitative brain electrical activity in the initial screening of mild traumatic brain injuries. *West J Emerg Med* 2012;13:394–400.
- McCrea M. Standardized mental status testing on the sideline after sport-related concussion. *J Athlet Train* 2001;36:274.
- McCrea M, Guskiewicz KM, Marshall SW, et al. Acute effects and recovery time following concussion in collegiate football players: the NCAA Concussion Study. *JAMA* 2003;290:2556–63.
- Randolph C, Millis S, Barr WB, et al. Concussion symptom inventory: an empirically derived scale for monitoring resolution of symptoms following sport-related concussion. *Arch Clin Neuropsychol* 2009;24:219–29.
- Prichep LS, Jacquin A, Filipenko J, et al. Classification of traumatic brain injury severity using informed data reduction in a series of binary classifier algorithms. *IEEE Trans Neural Syst Rehabil Eng* 2012;20:806–22.
- Prichep LS, McCrea M, Barr W, et al. Time course of clinical and electrophysiological recovery after sport-related concussion. *J Head Trauma Rehabil* 2013;28:266–73.

23. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B* 1996;267–88.
24. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging* 2004;17:205–16.
25. Linton O, Tenforde TS, Amis ES, et al. Summary of workshop on CT in emergency medicine: ensuring appropriate use. *J Am Coll Radiol* 2011;8:325–9.
26. O’Neil B, Naunheim R, DeLorenzo R. CT positive brain injury in mild TBI patients presenting with normal SAC scores. *Mil Med* 2014;179:1250–3.
27. Kerber KA, Schweigler L, West BT, et al. Value of computed tomography scans in ED dizziness visits: analysis from a nationally representative sample. *Am J Emerg Med* 2010;28:1030–6.
28. Melnick ER, Keegan J, Taylor RA. Redefining overuse to include costs: a decision analysis for computed tomography in minor head injury. *Jt Comm J Qual Patient Saf* 2015;41:313–22.
29. Newman-Toker DE, Edlow JA. High-stakes diagnostic decision rules for serious disorders: the Ottawa subarachnoid hemorrhage rule. *JAMA* 2013;310:1237–9.
30. Newman-Toker DE, McDonald KM, Meltzer DO, et al. How much diagnostic safety can we afford, and how should we decide? A health economics perspective. *BMJ Qual Saf* 2013;22 Suppl 2:ii11–20.