American Journal of Emergency Medicine 32 (2014) 606–608

Contents lists available at ScienceDirect

American Journal of Emergency Medicine

journal homepage: www.elsevier.com/locate/ajem

Original Contribution

Differences in interpretation of cranial computed tomography in ED traumatic brain injury patients by expert neuroradiologists☆

J. Stephen Huff, MD [a,*], Sandeep Jahar, DO, MS [b]

[a] Departments of Emergency Medicine and Neurology, University of Virginia, PO Box 800699, Charlottesville, VA 22908-0699
[b] Department of Emergency Medicine, UCONN School of Medicine, Farmington, CT

ARTICLE INFO

Article history:
Received 30 December 2013
Received in revised form 8 March 2014
Accepted 9 March 2014

ABSTRACT

Cranial computed tomography (CT) is generally regarded as the standard for evaluation of structural brain injury in patients with traumatic brain injury (TBI) presenting to the emergency department (ED). However, the subjective nature of the visual interpretations of CT scans and the qualitative nature of reporting may lead to poor interrater reliability. This is significant because CT positive scans include a continuum of structural injury with differences in treatment. The purpose of the present study was to evaluate the consistency of readings of head CT scans obtained within 24 hours after mild TBI in the ED, as assessed by an independent adjudication panel of 3 experienced neuroradiologists. In 80.1% of the cases, all 3 adjudicators agreed with the determination of the presence of structural injury. However, when interrater agreement was assessed with respect to the specific classification of the injury, agreement was poor, with a κ of 0.3 (0.29-0.316; confidence interval [CI] 95%). When classification was collapsed, considering only the presence or absence of hematomas, agreement among all 3 adjudicators improved to 55%, but the κ of 0.355, (0.332-0.78; CI 95%) was still only fair. The data suggest the need for improved recognition and quantification of specific structural injuries in the TBI population for better identification of patients requiring clinical intervention.

1. Introduction

Cranial computed tomography (CT) is generally regarded as the standard for the evaluation of structural brain injury in patients presenting to the emergency department (ED) with traumatic brain injury (TBI). The reliability of CT findings may be limited by the subjective nature of the visual interpretations and by the qualitative nature of reporting as well as the experience and training of the reader (eg, ED physician, radiologist, or neuroradiologist) [1-3]. This is further complicated by factors such as the period of elapsed time between injury and the CT scan and the exact nature of the brain injury [4]. Although there are limited studies that evaluate interrater reliability and intrarater replicability in the identification of TBI using CT, those that exist raise clinical concerns, even with expert raters. Molina et al [5] found up to a 50% failure rate in identifying different etiology of TBI on perimortem CT scans when compared with autopsy findings. In the evaluation of 100 suspected TBI patients, Laalo et al [2] compared CT of a consensus panel of 3 neuroradiologists with readings of 2 neuroradiologists, one neuroradiologist in training and the on-call staff radiologist. The on-call radiologist missed 67% of brain contusions, and although experience increased accuracy, there were marked differences noted even between the skilled readers. The authors conclude that, even between the reports of the most experienced readers, marked differences were seen.

Furthermore, CT positive (CT+) scans include a continuum of injury from those that are not life threatening ("clinically unimportant") to those that are potentially life threatening ("clinically important"), with associated differences in appropriate clinical response and action required [6]. It has long been recognized that the type of structural brain injury influences patient outcome, particularly because neuroimaging of patients with suspected TBI has become routine [7].

The purpose of the present study was to evaluate the consistency of readings of head CT scans obtained as part of the clinical evaluation in the ED after TBI, acquired as part of a large, prospective, validation study[1] and assessed by an independent adjudication panel of 3 experienced neuroradiologists. Interrater reliability with respect both to presence of structural injury as well as the etiology of the injury was assessed.

☆ Funding Source: This research was supported in part by BrainScope Company, Inc, Bethesda, MD, who granted expenses related to data acquisition and in part from the Department of Defense, contract no. W911QY-12-C-0004, *Assessment of Head Injury in the Emergency Department*.

* Corresponding author.
E-mail address: huff@virginia.edu (J.S. Huff).

[1] Data were gathered as part of the B-AHEADII validation trial funded in part by BrainScope Co, Inc, Bethesda MD, who granted expenses related to data acquisition and from the Department of Defense, contract no. W911QY-12-C-0004, Assessment of Head Injury in the Emergency Department.

http://dx.doi.org/10.1016/j.ajem.2014.03.010
0735-6757/

2. Methods

2.1. Patient population

Patients were enrolled at 11 medical centers within the United States, as part of a prospective clinical trial to detect structural brain injury in TBI. Patients considered for study were adults between the ages of 18 to 80 years, with Glasgow Coma Scale (GCS) of 9 to 15, who presented to the ED after a closed head injury and who had a CT scan of the head ordered as part of their standard clinical evaluation. The mean time interval between head injury and CT scan was 5:45, with 80% occurring within 8 hours of injury. Patients were excluded if they were obtunded to the point where they could not provide informed consent. In addition, those with advanced dementias, Parkinson disease, current chronic drug or alcohol dependence, known seizure disorder, tumors, history of brain surgery, mental retardation, pregnancy, or those taking daily prescribed medication for a psychiatric disorder were not eligible for the study. Approval for the study was obtained from all local institutional review boards. All subjects signed written informed consent except in one case who was consented by a surrogate.

2.2. CT adjudication panel

The radiologist assigned at the clinical enrollment site made initial cranial CT readings, and these were used as part of the patient's clinical disposition. Images were later electronically transmitted to a coordinating center using Digital Imaging and Communications in Medicine (DICOM) standards and overread blindly by 3 randomly chosen from a pool of 4 independent, board-certified neuroradiologists affiliated with the clinical research organization managing the study. The neuroradiologists were practicing at 2 major US teaching hospitals during the course of the study and had specific experience with acute head trauma. Adjudicators were formally trained on study methodology for review to reduce differences among them due to uncertainty about assessment criteria, (eg, definitions of classifications considered as evidence of structural brain injury in the traumatic TBI population), Marshall Scale (MS) scoring, assessment of basilar skull fracture, and criteria for non-TBI. The presence (CT+) or absence (CT negative [CT−]) of structure injury was determined by majority rule of this panel for the purposes of the study, and although classification was recorded, agreement in classification was not required in determining majority rule. The adjudicating neuroradiologists were supplied the patient's age and were blind to all other clinical or demographic information.

Analyses included calculation of the percent agreement across adjudicators as to the presence or absence of structural brain injury in the scan (ie, CT+ or CT−). A second analysis was related to the agreement within the CT+ scans as to the classification of the brain injury. The Table shows the detailed categories into which reports were classified. In addition, agreement for a more gross classification was conducted, where multiple categories were combined into those which included a hematoma (subdural hematoma (SDH) and/or epidural hematoma (EDH) and/or intracerebral hematoma (ICH)), those which did not include a hematoma but did include subarachnoid hemorrhage (SAH) and or contusions, and a third category for those which fell outside the other 2. Thus, the gross categories combined in the following from the Table, hematomas (category 1, 2, 3, 4, and 7), SAH and/or contusion (5 and 6), and other (8). Fleiss' κ was used to determine interrater reliability using MatLab 7.9.0 (The MathWorks, Inc., Natick, MA, USA), where 1 is perfect agreement, and 0 is chance.

Table
Detailed classification categories for CT+ determinations

Category	Description	% Of reads
1	SDH and/or EDH, with or without ICH	12.7%
2	No. 1 + SAH	21.7%
3	ICH only	4.9%
4	No. 3 + contusion and/or SAH	12.7%
5	SAH only	13.1%
6	Contusion only or w/ SAH	9.0%
7	Multiple hematomas + SAH and/or contusion	18.0%
8	Other (including CT−, diffuse axonal injury, cerebral swelling)	7.9%

3. Results

Of the 137 scans that were submitted as CT+ by the treating hospitals to the blinded panel of 3 neuroradiologists for independent review, 116 were classified as CT+ by protocol parameters. Thus, agreement between clinical sites and the independent adjudication panel on CT+ scans was 84.7%. Of the cases, where disagreement with the site read occurred 11 (52.4%) times, there was unanimous consensus of the adjudicators, that is, 3 of 3 agreed that the scan was negative (site had called positive). In the remaining cases, 2 of 3 of the adjudicators agreed that the CT was negative for acute intracranial injury.

The 116 patients adjudicated as CT+ included 84 males and 32 females, with a mean age of 45.17 (SD, 18.94). Eighty-two percent of the cases had GCS of 15, with the mean GCS of 14.64 (SD, 1.008). The MS scores for the vast majority (92%) was MSII, the most mild of the CT+ scores on the MS. Of the remaining cases, 6 were MS = III, 2 were MS = IV, and one was MS >IV. Thus, the structural injuries in the study population were heavily biased toward minor structural injury as rated by MS. All 3 adjudicators agreed on the CT+ determination, 93 (80.2%) of 116 of the time, and 2 of 3 agreed on the positive determination in the remaining 23 (19.9%) of 116 cases.

During the review of each scan, if the panel member determined that it was positive (CT+), the specific classification(s) of the injury observed was indicated. The data on classification were then placed into categories as described above. When detailed descriptions were taken into consideration, there was disagreement on the exact classification of the injury in 89 (19.9%) of the 116 cases. The Fleiss' κ for interrater reliability for detailed classification among the 3 adjudicators was found to be 0.3 (0.29-.316; confidence interval [CI] 95%). This indicates less than moderate agreement (0.41-0.60) [8].

Of the 89 cases, where all 3 of the adjudicators did not agree, 2 of the 3 agreed in 55 (62%) of them. There were 19 cases that had no agreement at all on specific classification across all 3. Of the remaining 15 cases that showed disagreement, only 2 of the 3 called the case positive, and of these cases, only 47% (7/15) agreed on specific classification. In general, most cases, where disagreement occurred involved more complex injuries with multiple types of traumatic injury, (eg, hemorrhages, hematomas, contusions, etc), whereas most cases, where agreement occurred were categorized as hematomas or SAH. Thus, disagreement was not limited to the least severe of the injuries.

When the categories are reduced to reflect only absence or presence of hematomas (the most severe and most clinically important in that clinical action is likely warranted), agreement was found to be 55% between all 3 adjudicators. The Fleiss' κ for interrater reliability for the grouped etiologies was found to be 0.355 (0.332-0.378; CI 95%), again below the level considered for moderate agreement.

4. Discussion

The current standard for imaging of acute TBI is based on human visual inspection of the CT scan without a form of quantification or measurement. Such high subjectivity leads to poor interrater reliability, especially when considering the severity of the injury. This is particularly significant because CT+ scans include a continuum of structural injury with differences in clinical response and action required. It is also noted that the adjudication panel in this study was composed of experienced neuroradiologists, whereas in the ED environment, the initial reading may not be done by a neuroradiologist potentially amplifying this problem.

Results of this study suggest that there is poor agreement on classification of injury in reading CT scans for patients presenting to the ED with closed head injury, even between expert neuroradiologists. Although this population presented with mild injury with high GCS and ability to participate in the study protocol, their structural injuries did include hematomas with the need for clinical interventions. Thus, the poor agreement on classification must be taken into consideration when integrating CT findings into the clinical treatment path. This suggests the need for improved quantification and classification of structural injury in the head injured population. With current interest in concussion, additional biomarkers other than CT are needed to define the process.

Acknowledgment

We thank the contributions of those who made this research possible, including the primary investigators and research technicians at all the clinical sites and the patients who volunteered. In addition, we acknowledge the editorial and statistical contribution of Leslie S. Prichep, PhD.

References

[1] Erly WK, Berger WG, Krupinski E, Seeger JF, Guisto JA. Radiology resident evaluation of head CT scan orders in the emergency department. AJNR Am J Neuroradiol 2002;23(1):103–7.
[2] Laalo JP, Kurki TJ, Sonninen PH, Tenovuo OS. Reliability of diagnosis of traumatic brain injury by computed tomography in the acute phase. J Neurotrauma 2009;26:2169–78.
[3] Khoo NC, Duffy M. “Out of hours” non-contrast head CT scan interpretation by senior emergency department medical staff. Emerg Med Australas 2007;19(2):122–8.
[4] Oertel M, Kelly DF, McArthur D, Boscardin WJ, Glenn TC, Lee JH, et al. Progressive hemorrhage after head trauma: predictors and consequences of the evolving injury. J Neurosurg 2002;96(1):109–16.
[5] Molina DK, Nichols JJ, DiMaio VJ. The sensitivity of computed tomography (CT) scans in detecting trauma: are CT scans reliable enough for courtroom testimony? J Trauma 2007;63:625–9.
[6] Stiell IG, Lesiuk H, Vandemheen K, Clement C, Reardon M, Eisenhauer MA, et al. Obtaining consensus for the definition of “clinically important” brain injury in the CCC study. Acad Emerg Med 2010;7(5):572.
[7] Gennarelli TA, Spielman GM, Langfitt TW, Gildenberg PL, Harrington T, Jane JA, et al. Influence of the type of intracranial lesion on outcome from severe head injury. J Neurosurg 1982;56(1):26–32.
[8] Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37(5):360–3.