Accuracy and Clinical Utility of Reports from Outside Hospitals for CT of the Cervical Spine in Blunt Trauma


ABSTRACT

BACKGROUND AND PURPOSE: Multidetector CT is the workhorse for detecting blunt cervical spine injury. There is no standard of care for re-interpretation of radiology images for patients with blunt trauma transferred to a higher level of care. The clinical impact of discrepancies of cervical spine CT reads remains unclear. We evaluated the discordance between primary (from referring hospitals) and secondary radiology interpretations (from a receiving level I tertiary trauma center) of cervical spine CT scans in patients with blunt trauma and assessed the clinical implications of missed cervical spine fractures.

MATERIALS AND METHODS: Medical records of patients with blunt trauma transferred to our institution between 2008 and 2015 were reviewed. Primary and secondary interpretations were compared and categorized as concordant and discordant. Two senior neuroradiologists adjudicated discordant reports. The benefit of re-interpretation was determined. For discordant cases, outcomes at discharge, injury severity pattern, treatment, and arrival in a cervical collar were assessed.

RESULTS: Six hundred fifty patients were included; 608 (94%) presented with concordant reports: 401 (61.7%) with fractures and 207 (31.8%) with no fractures. There were 42 (6.5%) discordant reports; 18 (2.8%) were cervical spine injuries undetected on the primary interpretation. Following adjudication, the secondary interpretation improved the sensitivity (99.3% versus 95.7%) and specificity (99.1% versus 91.7%) in detecting cervical spine fractures compared with the primary interpretation alone (P < .001).

CONCLUSIONS: There was an overall 6.5% discordance rate between primary and secondary interpretations of cervical spine CT scans. The secondary interpretation of the cervical spine CT increased the sensitivity and specificity of detecting cervical spine fractures in patients with blunt trauma transferred to higher-level care.

ABBREVIATIONS: CSI = cervical spine injury; ICD = International Classification of Diseases

Radiology studies are a frequent component of patient transfer between hospital emergency departments, especially with the increasing use of CT in the emergency department setting. Patients experiencing trauma are often transferred to higher levels of care on the basis of findings from screening imaging studies. The consequences of a missed traumatic injury can be serious, yet the clinical implications are scarcely discussed in the current blunt cervical spine trauma literature. According to the American College of Surgeons, level I and II trauma centers are required to have systems in place to view radiographic imaging from referring hospitals. Unfortunately, the absence of a standard of care in reviewing referring hospitals’ studies has led to important interinstitutional variability in handling these examinations, especially given differences in image-acquisition techniques, lack of previous comparison examinations, and variability in reimbursement.

A recent survey of radiologists and emergency providers at a large academic medical center revealed that while sharing images and reports from referring hospitals is highly valued, concerns regarding image quality and accurate interpretation remain. Although the lack of confidence may be related to academic bias, the literature offers support for secondary interpretations. A meta-analysis of discrepancy rates among primary and secondary interpretations across trauma and nontrauma imaging data showed an overall discrepancy rate of 32.2%, including a 20.4% discrepancy rate for major findings; the overall discrepancy rate for trauma studies was 19.7%. One possible reason may be that community hospitals encounter less frequent major trauma, and
variation in radiologists rotating through an emergency call pool may contribute to a higher miss rate than subspecialty-trained fellowship radiologists at a level I trauma academic center.\textsuperscript{14,15} To the authors’ knowledge, only 1 study has investigated the clinical impact of discrepancy in cervical spine CT interpretations.\textsuperscript{16} In their study, Khalilzadeh et al\textsuperscript{16} included only adult patients. In this population, they concluded that secondary reads improved diagnosis and benefited patient care.

The primary aim of this study was to evaluate the discordance between radiology interpretations from referring hospitals and secondary interpretations of cervical spine CT for patients with concern for cervical spine injury (CSI) transferred to our institution, a level I tertiary trauma center. A secondary aim was to evaluate the clinical implications of the discordant reports. We hypothesized that a secondary interpretation of referring hospitals’ cervical spine CTs would increase the detection of CSI, potentially impacting treatment and time to cervical spine clearance for patients experiencing trauma transferred to a higher level of care.

MATERIALS AND METHODS

Study Design and Data Collection

This retrospective study was approved by the University of Iowa institutional review board (IRB No. 201712728). Medical records of patients transferred to our institution, an academic level I trauma and tertiary care center, following blunt trauma injury between July 2008 and September 2015 were reviewed. Per our institutional protocol, all patients experiencing trauma accepted for transfer who have accompanying radiology images that are complete are requested to have a secondary interpretation of transferred radiology images. Inclusion criteria for data analysis were the availability of cervical spine CT images from the referring hospital with an external radiology report (primary interpretation) at the time of transfer and an available internal radiology interpretation (secondary interpretation) of the external images performed at admission. Patients were excluded if the referring hospital’s cervical spine images and reports were not available for review or if a secondary read was not performed at our institution. Patients were referred from 25+ surrounding regional/private nonacademic community hospitals in the United States, within a 200-mile radius of our institution. This study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines for cohort studies.\textsuperscript{17,18}

A thorough chart review was performed to compare the primary with the secondary interpretations of the same CT images of the cervical spine. Images were received as trauma protocol scans of the cervical spine, with available axial, sagittal, and coronal multiplanar reformations with soft-tissue and bone windows. These were transferred electronically from the referring hospital or via a CD and were uploaded to our institution’s PACS and interpreted by fellowship-trained neuroradiology or musculoskeletal radiology faculty. For the patients who met the inclusion criteria, the following variables were collected from electronic medical records: patients’ demographics (age and sex) and principal and secondary diagnoses with the respective International Classification of Diseases-10 (ICD-10) codes.

For all included patients, primary interpretations were compared against the secondary interpretations and categorized into concordant or discordant groups. Discordant cases were then adjudicated by 2 senior board-certified neuroradiologists (B.P. and J.M., both with 15+ years of practice). Both adjudicators reviewed the external CT images and the primary and secondary radiology reports to adjudicate the presence of CSI. The adjudicators were blinded to the radiologists who interpreted the primary and secondary reports and the facility at which the imaging and interpretation took place. Adjudicators had access to MR imaging performed at our institution in 1742 patients to classify discordant reports more accurately. STIR sequences were part of the protocol for MR imaging performed at our institution.

For patients with discordant reports, we collected the following additional clinical information: neurologic symptoms at presentation, outcome at discharge (alive versus deceased), CSI stability pattern, operative-versus-nonoperative treatment, treatment with a cervical collar, and patient arriving in a cervical spine collar at the time of transfer. The authors defined the following injuries as unstable CSI patterns: atlanto-occipital dislocation, anterior atlantoaxial dislocation, Jefferson (C1 burst) fracture, C2 dens fracture, hangman fracture, flexion teardrop fracture, extension teardrop fracture, bilateral facet dislocation, anterior subluxation, multilevel fracture, and fracture dislocation.\textsuperscript{19}

Sample Size Calculation

A statistically adequate sample size was estimated using the Cohen $\kappa$ statistic based on Khalilzadeh et al,\textsuperscript{16} who showed that 31% of patients presented with cervical spine injury on the initial interpretation. With an SD of 0.2, a minimum of 633 subjects would be required to observe a 92% agreement between the primary radiology interpretation and the secondary interpretation performed at our center.

Statistical Analysis

Statistical analysis was performed using SPSS 25 (IBM). Descriptive statistics were obtained from demographic and outcomes data. Variables were expressed as descriptive statistics including number, percentage, mean, and SD. Interrater reliability was established with the Cohen $\kappa$ statistic.

RESULTS

Study Population

A sample of 1251 patients who experienced blunt traumatic injury and were transferred from a referring hospital with a “store and interpret” order of the cervical spine were identified from July 1, 2008, to September 31, 2015. Of these patients, 601 patients (48%) did not have external images or a primary interpretation available and/or did not have a secondary interpretation of the available external images. These patients were excluded from the study. External images and primary interpretations as well as secondary interpretations were available for the 650 patients who were included in this study. As we reached the sample size determination, further data collection was stopped. Of the 650 patients who met the inclusion criteria, 374 (57%) were male. Ages ranged from 7 to 90 years (mean, 55 [SD, 23] years). Twenty of 650 patients were younger than 18 years of age (age
Concordant Interpretations

Of the 650 patients, 608 (94%) had a concordant interpretation when comparing the primary interpretation with the secondary interpretation. Four hundred one patients (62%) had CSI identified on both the primary and secondary interpretations and were classified as having concordant reports with CSI (Fig 1). The most common CSIs identified by ICD codes were closed fractures of the second cervical vertebra (11.4%), seventh cervical vertebra (5.2%), and first cervical vertebra (4.8%). Two hundred seven patients (32%) did not have a CSI on both the primary and secondary interpretations and were classified as having discordant reports without CSI.

Discordant Interpretations

There was a disagreement between the primary and secondary interpretations in 42 patients (6.5%), with only 1 discordant case involving a patient younger than 18 years of age. Adjudication of the discordant reports by 2 senior neuroradiologists determined that the primary interpretation correctly identified fractures in 3 of 42 reports (7.1%) and correctly identified no fractures in 2 reports (4.8%), with the secondary interpretation being incorrect. Adjudication of the discordant reports also determined that the secondary interpretation identified fractures in 18 of 42 reports (42.9%) that were missed on the primary interpretation and identified no fractures in 19 reports (45.2%) that were initially diagnosed in the primary interpretation. The results of adjudication revealed that 88.1% of the discordant reports matched the secondary interpretation; however, the secondary interpretation did not match the adjudication for 11.9% of discordant reports.

Compared with the primary interpretation alone, the secondary interpretation improved the sensitivity from 95.7% to 99.3% and specificity from 91.7% to 99.1% for detecting cervical spine fractures. Positive predictive values (95.5%–99.5%) and negative predictive values (92.1%–98.7%) improved with the addition of a secondary interpretation compared with only a primary interpretation, a statistically significant improvement (P < .001) (Table 1). Interrater reliability analysis resulted in a κ statistic of 0.8581.

Use of MR Imaging

MR imaging was performed within 24–48 hours of admission as ordered by the clinical team when indicated to further evaluate injury. Within the discordant group, a secondary interpretation was reported in all except 1 case before MR imaging was performed during hospitalization. There were 17 cases in the discordant group in which MR imaging was available to adjudicators to classify reports more accurately. From those 17 cases, there were 2 instances in which the MR imaging proved the primary interpretation, positive for CSI, to be correct, which was missed on the secondary interpretation. There were 6 instances in which the MR imaging confirmed a positive CSI on the secondary interpretation that was missed on the primary interpretation. There were 6 instances in which MR imaging excluded injury, confirming the negative finding on the secondary interpretation when the primary interpretation reported a CSI. There was 1 case in which the MR imaging excluded injury, confirming the negative findings on the primary interpretation when the secondary interpretation reported an injury. Finally, there were 2 cases in which the MR imaging interpretation did not comment on a fracture when the primary interpretation reported no injury and a secondary interpretation reported CSI.

We further evaluated the discordant groups because management of these patients was most likely to be impacted. The most frequent CSIs identified on the primary interpretation with no consensus on the secondary interpretation were facet fractures (32%, n = 6), vertebral body fractures (26%, n = 5), and lamina fractures (21%, n = 4) (Table 2). An example of a discordant falsely identified fracture is shown in Fig 2. The most frequent injuries undetected in the primary interpretation but detected in the secondary interpretation were fractures of the seventh cervical vertebra (transverse process, facet fractures: 50%, n = 9) and C2 injuries (lateral mass, dens, odontoid fractures: 28%, n = 5) (Table 2). In this population, 2 missed cervical spine injuries were classified as unstable (multilevel fractures at C2 and C5, and a C3 teardrop fracture) (Table 2). An example of a discordant missed injury on the primary interpretation, an unstable pattern teardrop fracture, is shown in Fig 3.

Clinical Implications

The subgroup of interest with respect to clinical outcomes was the discordant group in which secondary interpretation identified...
Table 2: Discordant primary interpretations with false-positives and false-negatives by frequencies of adjudicated injury types

<table>
<thead>
<tr>
<th>Discordant Reports</th>
<th>No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discordant falsely identified fractures</td>
<td>19/650</td>
<td>3</td>
</tr>
<tr>
<td>Discordant falsely identified fractures by type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facet</td>
<td>6/19</td>
<td>32</td>
</tr>
<tr>
<td>Vertebral body</td>
<td>5/19</td>
<td>26</td>
</tr>
<tr>
<td>Lamina</td>
<td>4/19</td>
<td>21</td>
</tr>
<tr>
<td>Transverse process</td>
<td>2/19</td>
<td>11</td>
</tr>
<tr>
<td>Spinous process</td>
<td>1/19</td>
<td>5</td>
</tr>
<tr>
<td>Subluxation</td>
<td>1/19</td>
<td>5</td>
</tr>
<tr>
<td>C1 arch</td>
<td>1/19</td>
<td>5</td>
</tr>
<tr>
<td>Discordant fractures missed on primary interpretation</td>
<td>18/650</td>
<td>2.8</td>
</tr>
<tr>
<td>Discordant missed fractures by type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7 (transverse process, facet, or endplate)</td>
<td>9/18</td>
<td>50</td>
</tr>
<tr>
<td>C2 (dens, odontoid, pedicle, transverse foramen, lateral mass)</td>
<td>5/18</td>
<td>28</td>
</tr>
<tr>
<td>C4 (spinous process, vertebral body, foramina)</td>
<td>4/18</td>
<td>22</td>
</tr>
<tr>
<td>Occipital condyle</td>
<td>2/18</td>
<td>11</td>
</tr>
<tr>
<td>C5 (transverse process, foramina)</td>
<td>2/18</td>
<td>11</td>
</tr>
<tr>
<td>Teardrop</td>
<td>1/18</td>
<td>5.5</td>
</tr>
<tr>
<td>C1 arch</td>
<td>1/18</td>
<td>5.5</td>
</tr>
</tbody>
</table>

*Data presented in this table are those of patients with discordant results postadjudication. Note that some cases had multiple cervical spinal fractures.

FIG 2. Example of a falsely diagnosed fracture on primary interpretation. The referring hospital falsely identified a chip fracture of the right inferior articular facet of C5 (CT images, A and C). MR imaging was performed to assess right upper extremity weakness and did not show any evidence of fracture or ligamentous injury. Axial T1 (B) and sagittal STIR (D).

an injury not identified in the primary interpretation, 2 of which were adjudicated as no fracture present and 18 with fracture present (Online Supplemental Data). Two of the 18 patients had an unstable CSI pattern at the secondary interpretation; 1 of those 2 patients required an operation. The unstable patterns not appreciated on the primary interpretation included multilevel C2 and C5 fractures and a C3 teardrop fracture (Fig 3). Both patients with the missed unstable injury pattern (teardrop fracture and C2 and C5 fractures) had a clinical diagnosis of central cord syndrome based on physical examination. Two patients, both found to have stable injury patterns on the secondary interpretation and subsequently managed with cervical immobilization, arrived from referring hospitals without a cervical collar. The remaining 16 of 18 patients with missed CSIs on the primary interpretation were treated with cervical immobilization. There was 1 mortality in this group, but it was unrelatable to the cervical spine pathology.

DISCUSSION

Access to accurate radiographic interpretations is paramount for providers. In the case of transferred patients, treating providers may be biased to radiographic interpretations from their own facility. The decision to obtain secondary interpretations of imaging performed from referring hospitals should be considered on the basis of current literature evidence and local assessments of discordance rates, considering time, resources, and the cost of overreads or missed injuries. Sung et al identified a 12% discordance rate for outside trauma imaging requested for secondary interpretation at higher level of care centers; their scans were inclusive of abdomen/pelvis, brain, chest, and other body region studies. More recently, Robinson et al showed an 8.5% rate of discrepancy between referring hospitals’ primary interpretations and internal secondary interpretations in the emergency department for imaging of transfer patients with and without trauma. Because surgical subspecialists, including spine specialists, are not always available in our catchment area, a large number of patients are transferred for management of injuries identified on imaging.

Our center performs secondary interpretations as ordered by the treating providers. The purpose of this study was to examine the discrepancy rates specifically in cervical spine imaging. Our goals were the following: 1) add to the relatively sparse body of literature available for institutions considering secondary interpretation of trauma imaging, 2) identify patterns in the missed and falsely diagnosed cervical spine fractures, 3) evaluate the clinical implications of a missed CSI, and 4) assess whether the practice of secondary interpretations of transferred imaging should continue.

We found a 6.5% rate of discordance in cervical spine CT interpretations in our study. This is similar to findings of Khalilzadeh et al, who found an 8% discordance rate between referring hospital CT cervical spine reports and internal secondary interpretations in patients experiencing trauma transferred to a higher level of care. This is the only additional study in the literature to look specifically at discordance rates in cervical spine imaging in patients experiencing trauma. The sample size of our study is double that of Khalilzadeh et al. While there are studies that cite average generalized diagnostic error rates, this study focused specifically on the subset of cervical spine CT imaging in patients experiencing trauma, comparing concordance and discordance between radiology
providers covering rural community hospitals and radiologists at an academic level I trauma center to determine whether there is utility in requesting a secondary interpretation on patient transfer. This study adds to the literature because it presents the largest cohort studied to date. It is the first to include pediatric patients and to report the clinical implications of a missed CSI, including management of injury, operative repair, complications pertaining to CSI, and investigation of further advanced imaging. The methodology of this study including the adjudication process is a strength compared with current literature.

For all 42 discordant cases, all primary reads were from community-based radiologists. Although the authors are not privy to the training background of the primary radiologists, we hypothesize that the discordance rate and missed CSI on primary interpretation is explained by the variability of general radiologists rotating through emergency calls (as opposed to dedicated emergency-trained radiologists), with comparatively less frequent trauma exposure in smaller community hospitals. The most frequently missed radiologic findings in this study population were transverse process fractures and dens fractures, a pattern consistent with the prior study. We observed a relatively higher frequency of missed C7 fractures in this population, for which there is not a strong explanation. External cervical spine CTs were performed with a trauma protocol and imaged through T2–T3 vertebral bodies on the sagittal view, providing for complete examinations. Satisfaction of search may explain missed C7 injuries because this region of the cervical spine often falls at the end of the search pattern. The most frequent overcalled findings were facet fractures and lamina fractures. In addition, the miss rate of CSI can be influenced by image quality, the presence of multilevel degenerative disease, existing spine instrumentation/hardware, and normal variants (accessory ossicles, limbus vertebrae, unfused apophysis), which add to the challenge of interpreting trauma CT scans of the cervical spine and its instrumentation.

FIG 3. Example of a missed, unstable pattern fracture on the primary interpretation from the referring hospital, which was identified as a C3 teardrop fracture seen in both the coronal (left) and sagittal (right) reformation images on the secondary interpretation (arrows). The patient presented with upper extremity weakness and was treated nonoperatively with a cervical collar.

The clinical implications of missing cervical spine injuries should be an important factor in the decision to perform secondary interpretations of imaging in transferred patients. In this study population, 2 of 18 patients with missed fractures on primary interpretations had an unstable fracture pattern. Another 2 of 20 were transferred to a higher level of care without cervical immobilization. Two patients of 18 presented with neurologic symptoms. Although an overall small percentage of patients fell into this group, the authors favored the secondary interpretation because of the improved sensitivity and specificity of detecting cervical spine fractures following secondary interpretation. The authors recognize that very few discrepancies led to surgical interventions or were unstable fractures affecting morbidity or mortality. The 6.5% discrepancy rate may, in part, reflect an underlying 3%–5% diagnostic error rate reported in the literature. However, this percentage is a general estimate that includes multiple imaging modalities, whereas our study evaluates a specific imaging format (cervical spine CT) in a specific population (patients with blunt trauma) between community hospitals in rural areas that see trauma less frequently and a level I trauma academic center. The decision to recommend a secondary interpretation as a standard of practice will depend on locoregional error rates, imaging technique, and acuity of the injury, among other factors such as reimbursement. Previous studies supporting secondary interpretations of imaging performed in transferred patients have found decreased rates of subsequent imaging use and improved patient diagnosis and care. Arguments against secondary interpretation would be a lack of standardized protocols in liability and reimbursement for external interpretations, added time and resource use, and clinical delay of clearing negative findings on cervical spine examinations for patients experiencing trauma.

This study has some limitations. First, this is a retrospective study that introduces a risk of selection bias and incomplete data capture from poor documentation or missing data. From the 1251 patients retrospectively identified, 48% did not have external radiology imaging or a primary interpretation on transfer to our institution. For some transfers, cervical spine imaging was not performed at the referring hospital, depending on the type of traumatic injury. Subsequently, transfer patients in whom cervical spine imaging was performed on the basis of clinical suspicion or mechanism of injury resulted in a higher rate of CSI detection (61.7%) compared with the literature. This is likely because this study specifically looked at patients transferred for concern for spine injury in the setting of blunt trauma. Although our study included pediatric patients,
only 20 of the 650 included cases were pediatric cases, and only 1 discordant case involved a child, limiting our impact to comment on discordant rates in pediatric populations.

The study did not address additional fractures that were missed in the cervical spine, and discordance was based on whether a cervical spine injury was called. The smaller sample size of discordant reports also limited our ability to perform a more robust statistical analysis pertaining to clinical outcomes. Future studies comparing outcomes between the discordant and concordant patient groups are warranted. The authors also acknowledge that the imaging quality, windowing, and motion of the external hospital CT scanner were factors beyond the scope and control of this study and may have impacted image interpretation. Finally, our demographic analysis did not include identification of the mechanism of injury, which could affect the odds of having CSI.

CONCLUSIONS

In this review of 650 cervical spine CT studies on patients transferred from community hospitals to a level I trauma center, there was a 6.5% discordance rate between the primary and secondary interpretations of the CT scans of the cervical spine, with 3% of patients presenting with a CSI not appreciated on the primary interpretation. This discordance impacted both the nonoperative and surgical management of patients in this study. The secondary interpretation significantly improved the sensitivity and specificity for the detection of cervical spine fractures in patients transferred to a higher-level care. While re-interpretating external CT scans of patients with blunt trauma poses challenges with respect to time, resource allocation, medicolegal liability, and compensation, secondary interpretations of external cervical spine CTs present improved detection of injuries resulting in a benefit in terms of care and safety for patients with blunt trauma transferred to a higher level of care. On the basis of the findings of this study and the existing literature, the authors suggest that secondary interpretation of CT scans for patients with blunt trauma transferred to a higher level of care should become a standard practice nationally, but they recognize that the decision to do so is additionally influenced by the baseline locoregional diagnostic error rate and policies in reimbursement.

ACKNOWLEDGMENTS

The authors would like to acknowledge Tzu Chuan Yen for her contribution in chart review.

Disclosures: Jacklyn Engelbart—UNRELATED: Grants/Grants Pending: National Institutes of Health. Comments: I am a research resident with my salary funded by the National Institutes of Health T32 Surgical Oncology Research Grant at my institution. This project was completed outside of the scope of that research and funding.* *Money paid to the institution.

REFERENCES

6. American College of Surgeons, Committee on Trauma. Resources for Optimal Care of the Injured Patient. 6th ed. America College of Surgeons, Committee on Trauma; 2014

