AJNR

CT Cervical Spine Fracture Detection Using a Convolutional Neural Network

J.E. Small, P. Osler, A.B. Paul and M. Kunst

AJNR Am J Neuroradiol 2021, 42 (7) 1341-1347
doi: https://doi.org/10.3174/ajnr.A7094
http://www.ajnr.org/content/42/7/1341

CT Cervical Spine Fracture Detection Using a Convolutional Neural Network

J.E. Small, P. Osler, A.B. Paul, and M. Kunst

ABSTRACT

BACKGROUND AND PURPOSE: Multidetector CT has emerged as the standard of care imaging technique to evaluate cervical spine trauma. Our aim was to evaluate the performance of a convolutional neural network in the detection of cervical spine fractures on CT.

MATERIALS AND METHODS: We evaluated C-spine, an FDA-approved convolutional neural network developed by Aidoc to detect cervical spine fractures on CT. A total of 665 examinations were included in our analysis. Ground truth was established by retrospective visualization of a fracture on CT by using all available CT, MR imaging, and convolutional neural network output information. The κ coefficients, sensitivity, specificity, and positive and negative predictive values were calculated with 95% CIs comparing diagnostic accuracy and agreement of the convolutional neural network and radiologist ratings, respectively, compared with ground truth.

RESULTS: Convolutional neural network accuracy in cervical spine fracture detection was 92% (95% CI, 90%–94%), with 76% (95% CI, 68%–83%) sensitivity and 97% (95% CI, 95%–98%) specificity. The radiologist accuracy was 95% (95% CI, 94%–97%), with 93% (95% CI, 88%–97%) sensitivity and 96% (95% CI, 94%–98%) specificity. Fractures missed by the convolutional neural network and by radiologists were similar by level and location and included fractured anterior osteophytes, transverse processes, and spinous processes, as well as lower cervical spine fractures that are often obscured by CT beam attenuation.

CONCLUSIONS: The convolutional neural network holds promise at both worklist prioritization and assisting radiologists in cervical spine fracture detection on CT. Understanding the strengths and weaknesses of the convolutional neural network is essential before its successful incorporation into clinical practice. Further refinements in sensitivity will improve convolutional neural network diagnostic utility.

ABBREVIATIONS: AI = artificial intelligence; CNN = convolutional neural network; NPV = negative predictive value; PPV = positive predictive value

A variety of studies have been conducted evaluating the performance of artificial intelligence (AI) to detect fractures. AI has been used to detect hip,[1-3] humeral,[4] distal radius,[5] wrist,[6-8] hand,[8] and ankle fractures[8] on radiographs, as well as thoracic and lumbar spine fractures on dual x-ray absorptiometry.[9] In addition, AI has been used to detect calcaneal[10] and thoracic and lumbar vertebral body fractures[11-13] on CT. To our knowledge, no studies evaluating AI in detecting cervical spine fractures on CT have been published.

Cervical spine injury is common with greater than 3 million patients per year being evaluated for cervical spine injury in North America,[14] and greater than 1 million patients with blunt trauma with suspected cervical spine injury per year being evaluated in the United States.[15] Cervical spine injury can be associated with high morbidity and mortality,[16] and a delay in diagnosis of an unstable fracture leading to inadequate immobilization may result in a catastrophic decline in neurologic function with devastating consequences.[17-20] Clearing the cervical spine through imaging is therefore a critical first step in the evaluation of patients with trauma, and multidetector CT has emerged as the standard of care imaging technique to evaluate cervical spine trauma.[21] Morbidity and mortality in patients with cervical spine injury can be reduced through rapid diagnosis and intervention.

The aim of this study is to evaluate the performance of a convolutional neural network (CNN) developed by Aidoc (www.aidoc.com) for the detection of cervical spine fractures on CT. We establish the presence of fractures based on retrospective clinical diagnosis and compare the CNN performance with that of radiologists. Aidoc's CNN currently runs continuously on our

Received August 21, 2020; accepted after revision January 25, 2021.

From the Departments of Neuroradiology (J.E.S., A.B.P., M.K.) and Radiology (P.O.), Lahey Hospital and Medical Center, Burlington, Massachusetts.

Paper previously presented as a poster at: Annual Meeting of the American Society of Neuroradiology, May 30 to June 4, 2020; Virtual.

Please address correspondence to Juan E. Small, MD, Lahey Hospital and Medical Center, 41 Mall Rd, Burlington, MA 01805; e-mail: Juan.E.Small@Lahey.org
http://dx.doi.org/10.3174/ajnr.A7094

hospital system and functions as a triage and notification software for analysis and detection of cervical spine fractures. However, we purposefully conducted a retrospective study on cervical spine studies performed before system-wide deployment, as we wanted to compare CNN performance to radiologist performance without the aid of the tool. A proficient algorithm may help identify and triage studies for the radiologist to review more urgently, helping to ensure faster diagnoses.

MATERIALS AND METHODS

After approval by our institutional review board, we conducted a retrospective analysis of the predictions of an FDA-approved CNN developed by Aidoc for the identification of cervical spine fractures based on CT. We compared these predictions to the unaided diagnoses made by radiologists of different levels of expertise and training. Our criterion standard for the presence or absence of cervical spine fractures was based on retrospective consensus review by 2 fellowship-trained neuroradiologists after evaluating all available CT, MR imaging, and CNN data.

CNN Algorithm Development

We evaluated an FDA-approved CNN developed by Aidoc for cervical spine fracture detection on CT. The CNN is designed to detect linear bony lucency in patterns consistent with fracture (including compression), does not distinguish between acute and chronic fractures, and is limited to the cervical spine (C1–7).

The hardware used for developing and validating the CNN included 8 GPUs, 64 CPUs, 488 GB of RAM, and 128 GB of GPU memory. Validation was based on retrospective, blinded data from 47 clinical sites evaluating approximately 8000 examinations. Nearly equal amounts of positive and negative examinations were included in the analysis. Validation sensitivity was 95.8% (95% CI, 95.7%–95.9%) and specificity was 98.5% (95% CI, 98.5%–98.5%). Approximately 12,000 studies from 83 clinical sites were used for training the algorithm, and 80% of them were positive. The CNN training data base was made from datasets from all commercially available CT scanners, and included all available imaging planes (axial, coronal, and sagittal) and kernels (bone and soft tissue). The training data base labeling was based on manual review and annotation of fractures by neuroradiologists experienced in spine trauma.

The cervical spine fracture detection model consists of 2 stages: a region proposal stage and a false-positive reduction stage. The first stage is a 3D fully convolutional deep neural network. The architecture is based on the Residual Network architecture, which consists of repeated blocks of several convolutional layers with skip connections between them, and is followed by a pooling layer that reduces the dimensions of the output. This network is trained on segmented scans and produces a 3D segmentation map. The model was trained from scratch, with no pretraining from additional datasets. From the segmentation map, region proposals are extracted and passed as input to the second stage of the algorithm. The second stage classifies each region as positive or negative. Two sets of features are extracted from each region, fused together, and used for the final decision. The first are learned features from a multilayered, classification head that receives the features from the last layer of the 3D segmentation network for the proposed regions as input. The second are nonlearned engineered features obtained from traditional image-processing methods that operate on the proposed regions. These features are combined through an additional neural network, which classifies each proposal as a fracture or not.

Validation Dataset

We queried the PACS for cervical spine CT studies performed between January 3, 2015, and December 30, 2018 (a time before system-wide deployment of the CNN algorithm at our institution), in patients who also had a short interval follow-up cervical spine MR imaging (<48 hours). In particular, we limited the analysis to cervical spine CT studies with a short interval follow-up MR imaging so that the MR imaging data could aid in the retrospective criterion standard determination of acute fractures. Of note, examinations at our institution have cervical spine MR imaging after a CT when there is a persistent clinical concern for cervical spine trauma despite a negative cervical spine CT, or in patients with positive cervical spine CT findings for trauma to evaluate for cord contusion, ligamentous injury, or epidural hemorrhage.

The CNN validation was made of datasets acquired from multiple institutions on all commercially available CT scanners with differences in FOV and section thickness. Similarly, the study group included datasets acquired on different commercially available CT scanners at both Lahey Hospital and Medical Center and affiliate institutions with differences in FOV and section thickness. MR images used to troubleshoot examinations were performed at both 1.5T and 3T and were not evaluated by the CNN. The finalized cervical spine CT reports were simultaneously independently reviewed by 2 fellowship-trained neuroradiologists. To achieve labeling consensus maximizing ground truth assessment in our study, the decision was made to have 2 neuroradiologists who had each completed a 2-year neuroradiology fellowship and obtained the Certificate of Added Qualification review each report. Results were classified as positive or negative for fracture.

Error Analysis

The cervical spine CTs were interpreted and dictated at the time of patient presentation by a diverse group of radiologists. This group consisted of neuroradiologists (some of whom had obtained the Certificate of Added Qualification), emergency department radiologists, general private practice radiologists from affiliate hospitals, and remote overnight coverage nighthawk radiologists (some of whom had completed fellowship training in neuroradiology). Meaningful analysis and conclusions comparing the CNN to different-level radiologists was not feasible because of the wide variety of training backgrounds and small number of radiologists within some of the groups. Research data analysis was performed by neuroradiologists who had completed a 2-year neuroradiology fellowship and obtained the Certificate of Added Qualification. Ground truth labeling was obtained by retrospective visualization of a fracture on CT after using all available CT, MR imaging, and CNN information and was performed independently by 2 fellowship-trained neuroradiologists. Discrepant CNN positive

examinations were reviewed both on a custom web-based viewer and in the PACS. The finalized CT reports and primary CNN output were graded against the ground truth.

Only fractures involving the cervical spine (C1–7), as well as both acute and chronic fractures, were labeled true-positives to match the design of the CNN. Postsurgical changes, congenital fusion anomalies, nutrient foramina, degenerative changes, and artifact were labeled negative for fracture. Traumatic disc injuries were labeled true-negatives as they do not match the design of the CNN by failing to contain a linear bony lucency in a pattern consistent with fracture.

Our study data base included datasets from several referring institutions with different scanner manufacturers and techniques mimicking the heterogeneity of the CNN training data base. Most of the examinations at our institution were performed on an Ingenuity CT scanner (Philips Healthcare) with 1.5 mm axial section thickness and 1 mm coronal and sagittal reformats.

Diagnostic accuracy and agreement between the radiologist and the ground truth, and between the CNN and ground truth, was evaluated by using κ coefficients, sensitivity/specificity, positive predictive value (PPV), and negative predictive value (NPV). The 95% confidence intervals were calculated for each estimate.

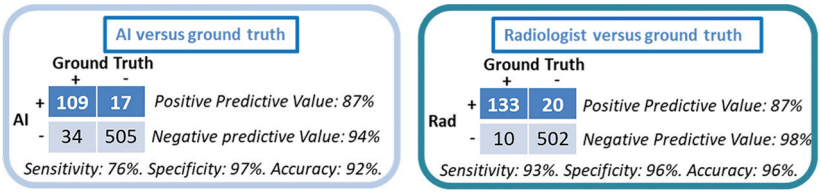

FIG 1. CNN and radiologist performance.

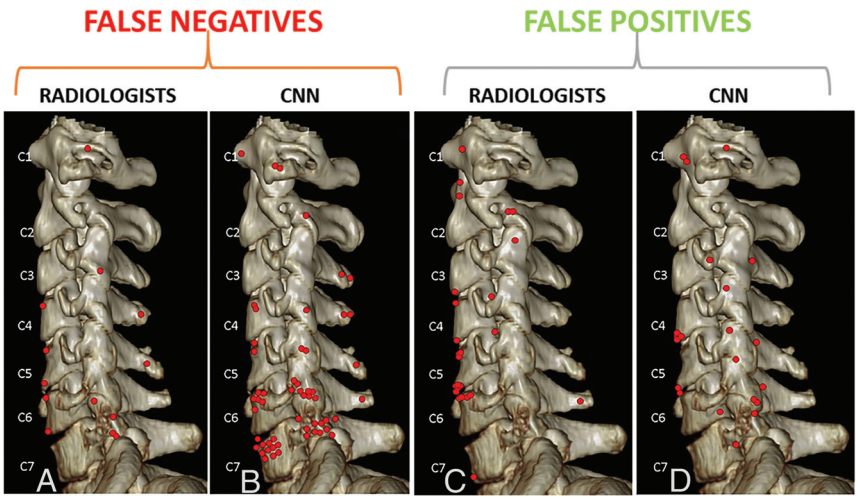

FIG 2. Location of false-negative and false-positive fractures for the radiologists and the CNN. Each location instance is marked by a red dot. False-negative fractures by radiologists (*A*) and the CNN (*B*) were similar in site and distribution. Although both the radiologists and CNN missed fractures more commonly along the lower cervical spine, errors were more numerous for the CNN. False-positive fracture sites noted by radiologists (*C*) and the CNN (*D*) can also be compared side-by-side. Numerous findings can mimic fractures on CT, most commonly degenerative changes or nutrient foramina. These fracture mimics were misinterpreted by both radiologists and the CNN. False-positive fractures along the anterior corners of vertebral bodies were slightly more commonly noted by radiologists and false-positive fractures along the facets and transverse processes were more commonly identified by the CNN.

RESULTS

A total of 869 cervical spine CT examinations were initially identified. The patient age range was 16–98 years, with an average of 60.28 years and a median of 61 years. A total of 379 patients were men (54.5%) and 316

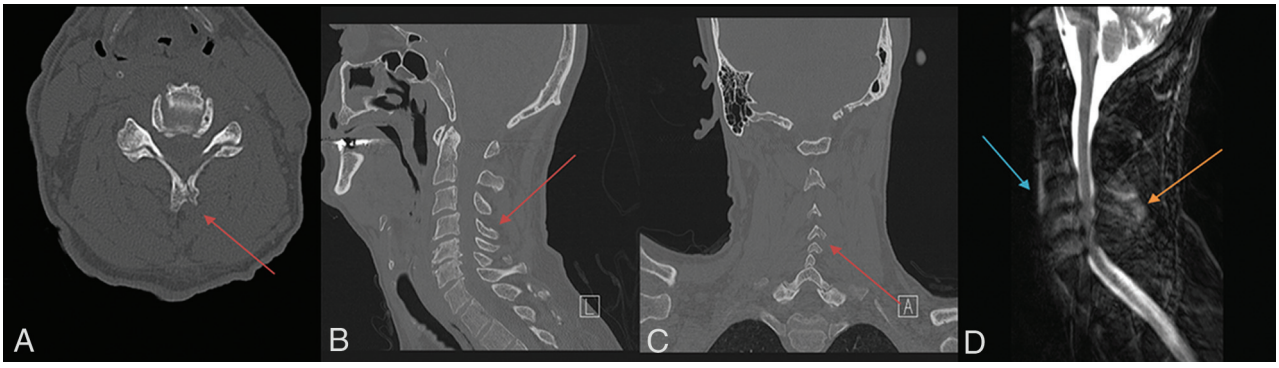

FIG 3. Fracture-positive, radiologist false-negative, CNN false-negative case example. Axial (*A*), sagittal (*B*), and coronal (*C*) cervical spine CT images, and sagittal fat-saturated T2-weighted cervical spine MR image (*D*) demonstrate a minimally displaced C4 spinous process fracture. *Red arrows* demarcate fracture lines, the *blue arrow* demarcates prevertebral edema, and the *orange arrow* demarcates interspinous and supraspinous ligamentous injury. This case example illustrates a subtle fracture missed by both the radiologist and CNN that was identified in retrospect with the help of MR imaging because of the presence of secondary signs, such a prevertebral edema and ligamentous injury.

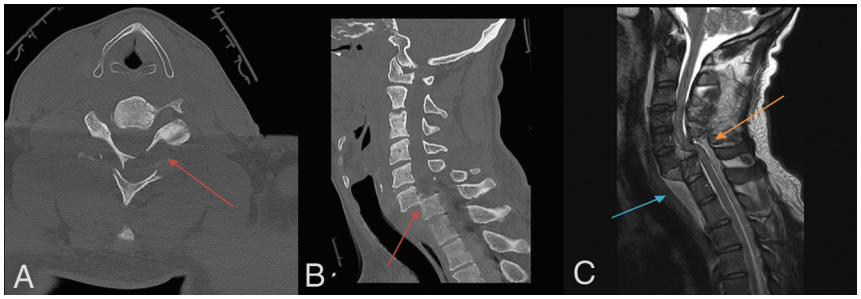

FIG 4. Fracture-positive, radiologist true-positive, CNN false-negative case example. Axial (*A*) and sagittal (*B*) cervical spine CT images, and sagittal fat-saturated T2-weighted cervical spine MR image (*C*) demonstrate a C6–7 fracture-dislocation with cord compression. *Red arrows* demarcate fracture-dislocation, the *blue arrow* demarcates prevertebral edema, and the *orange arrow* demarcates cord compression. This case example illustrates an important drawback of the CNN to overlook areas of gross bony translation, as it was only designed to detect linear bony lucency in patterns consistent with fractures.

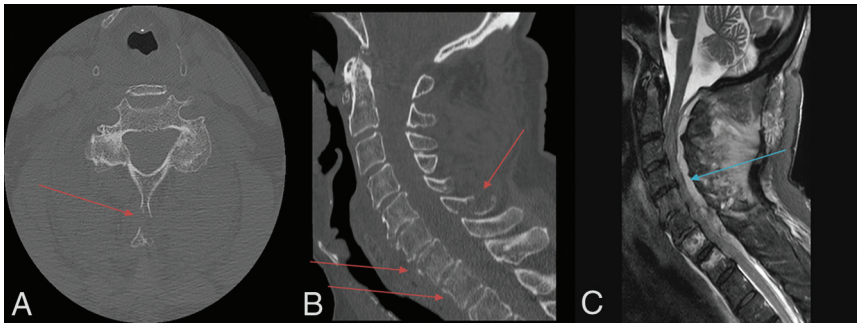

FIG 5. Fracture-positive, radiologist true-positive, CNN false-negative case example. Axial (*A*) and sagittal (*B*) cervical spine CT images, and sagittal fat-saturated T2-weighted cervical spine MR image (*C*) demonstrate multiple fractures involving the C7 and T1 vertebral bodies and C6 spinous process with epidural hematoma and associated cord compression. *Red arrows* demarcate fracture lines and the *blue arrow* demarcates epidural hematoma. This case example illustrates important drawbacks of the algorithm to miss fractures characterized more by distraction rather than linear bony lucency, fractures involving the distal aspects of the spinous processes that may be mistaken for nuchal ligament calcification or ossification, and fractures located in the lower cervical spine where fine bony detail becomes poor from CT beam attenuation.

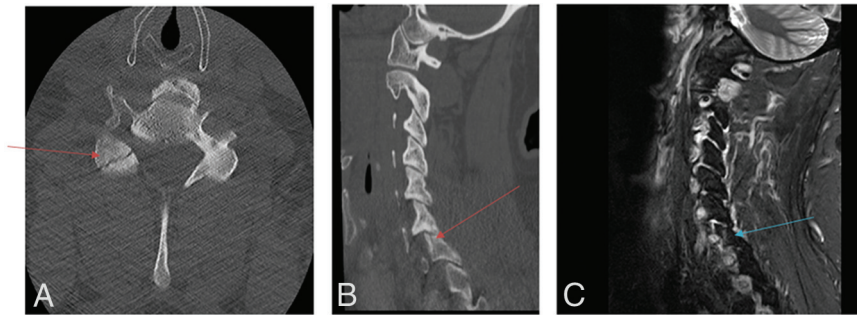

FIG 6. Fracture-positive, radiologist false negative, CNN true positive case example. Axial (*A*) and sagittal (*B*) cervical spine CT images and sagittal (*C*) STIR cervical spine MR image demonstrate a minimally displaced right C7 superior articulating facet fracture. *Red* and *blue arrows* demarcate fracture lines. This case example illustrates the strength of the algorithm to detect subtle fractures when they conform to linear bony lucencies.

were women (45.5%). Twelve examinations were duplicates and 162 examinations could not be processed by the CNN. A total of 157 of the 162 excluded examinations could not be retrieved from the PACS by the CNN because they were imported from an outside hospital without an identifiable DICOM header. The remaining 5 of the 162 excluded examinations could not be analyzed by the CNN because of technical issues with the datasets. These technical issues related to a few preprocessing steps of the CNN orchestrator, which assure that the study is technically adequate for analysis. These include inconsistent DICOM tags or missing slices that would compromise the processing. The fracture prevalence in the excluded dataset is similar to the fracture prevalence in the included dataset. For example, 35 of the 162 excluded examinations were positive for fracture (22%) compared with 143 of the 695 included examinations (21%). Because this was a retrospective study of datasets acquired before CNN implementation, the percentage of excluded examinations on datasets acquired after CNN implementation is likely to be much smaller based on the availability of technical support from the CNN developer and the presence of reliable DICOM tags. Consequently, we feel the true accuracy of the CNN to be comparable with the accuracy demonstrated in our study. Out of the 695 remaining examinations, 30 examinations had fractures outside of the cervical spine (C1–7) and were excluded from our analysis, for a final sample size of 665 examinations. A total of 143 examinations were labeled positive for fracture and 522 examinations were labeled negative for fracture by ground truth analysis.

For the radiologists, there were 133 examinations labeled true-positive in which fractures were noted in the report and 502 examinations labeled true-negative in which no fractures were noted in the report. There were 20 examinations labeled false-positive in which a fracture was mentioned in the report but both MR imaging and CNN output were negative for fracture. There were 10 examinations labeled false-negative in which no fracture was

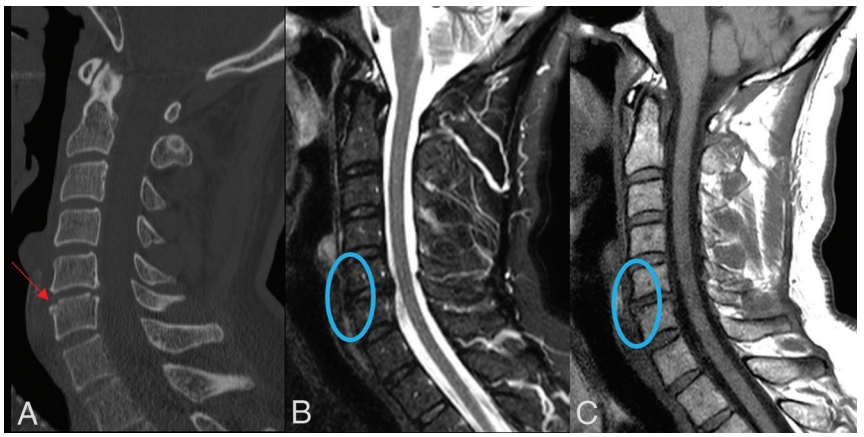

FIG 7. Fracture-negative, radiologist false positive, CNN false positive case example. Sagittal (*A*) cervical spine CT image and sagittal (*B*) STIR and sagittal (*C*) T1-weighted cervical spine MR images demonstrate a small depression along the anterosuperior margin of the C6 vertebral body (*red arrow*) without associated bone marrow edema, prevertebral edema, or disc space widening (*blue circles*). This case example illustrates how both the radiologist and CNN algorithm are capable of ignoring the absence of secondary signs, such as prevertebral edema and disc space widening when incorrectly identifying a fracture.

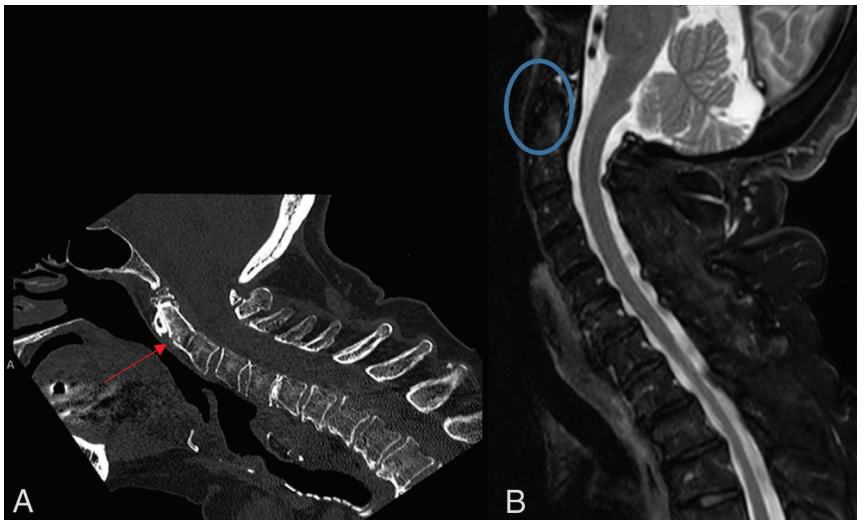

FIG 8. Fracture-negative, radiologist false positive, CNN true negative case example. Sagittal (*A*) cervical spine CT image and sagittal (*B*) STIR cervical spine MR image demonstrate a nutrient foramen/trabecular variant within the dens (*red arrow*) without associated bone marrow edema (*blue circle*). This case example illustrates the ability of the AI algorithm to exclude those linear bony lucencies that contain sclerotic margins inconsistent with fracture.

mentioned in the report but either MR imaging or CNN output were positive for fracture and the fracture could be visualized in retrospect on the cervical spine CT. The PPV and NPV for the radiologist was 87% (95% CI, 81%–92%) and 98% (95% CI, 96%–99%), respectively. The sensitivity, specificity, and percent agreement were 93% (95% CI, 88%–97%), 96% (95% CI, 94%–98%), and 95.5% (95% CI, 94%–97%), respectively. The κ coefficient was 0.87 (95% CI, 0.82–0.92). The time from acquisition until a finalized report for the radiologist ranged from 33 to 43 minutes.

For the CNN, there were 109 examinations labeled true-positive and 505 examinations labeled true-negative that matched ground truth labeling. There were 17 examinations labeled false-positive in which the CNN detected a fracture, but both the radiologist and MR imaging reports were negative for fracture. There were 34 examinations labeled false-negative in which the CNN failed to detect a fracture that was seen in both the radiologist and MR imaging reports. The PPVs and NPVs for the CNN were 87% (95% CI, 79%–92%) and 94% (95% CI, 91%–96%), respectively. The sensitivity, specificity, and percent agreement for the CNN was 76% (95% CI, 68%–83%), 97% (95%–98%), and 92% (95% CI, 90%–94%), respectively. The κ coefficient was 0.76 (95% CI, 0.70–0.82). The time from acquisition until a CNN analysis report ranged from 3 to 8 minutes.

To address the concern of selection bias in our sample (with an incidence of 21.5%), extrapolation to a population with an incidence of 1.9% as reported by Inaba et al[16] of cervical fracture was conducted. With the same sample size and values of sensitivity and specificity as found above, estimated PPVs and NPVs for the radiologist's ratings were 32% (95% CI, 18%–50%) and 99.9% (95% CI, 99%–100%), and for the CNN's ratings, they were 30% (95% CI, 15%–49%) and 99.5% (95% CI, 99%–100%).

In 7 examinations labeled true-positive, the CNN detected a fracture that the radiologist missed on CT and MR imaging. In 4 examinations labeled true-positive, the fracture detected by the CNN was chronic.

The results for CNN-versus-radiologist performance are summarized in Fig 1. The location of false-negative and false-positive fractures for the CNN and radiologist are compared in Fig 2. Several instructive examples are depicted in Figs 3–10.

DISCUSSION

We evaluated the performance of a CNN designed to detect cervical spine fractures on CT and compared it to that of radiologists. Our dataset contained a high fracture prevalence because of our decision to limit our analysis to only those examinations that

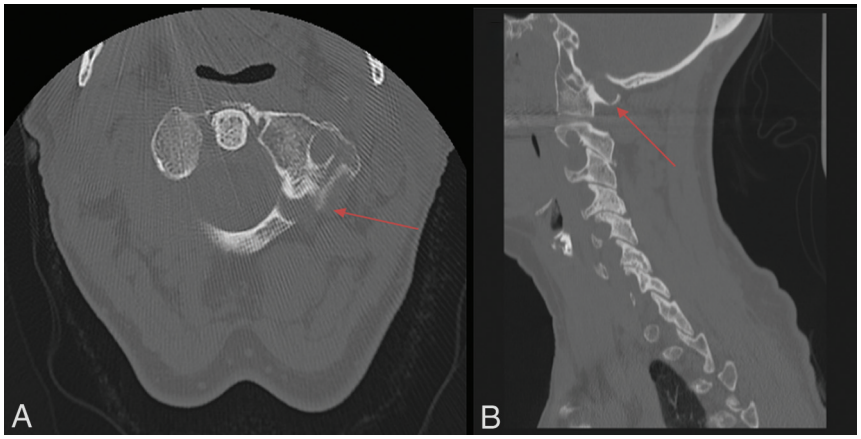

FIG 9. Fracture-negative, radiologist true negative, CNN false positive case example. Axial (*A*) and sagittal (*B*) cervical spine CT images demonstrate congenital thinning and incomplete fusion of the left C1 lamina. *Red arrows* demarcate congenital thinning and incomplete fusion. This case example illustrates a limitation of the AI algorithm to mistake common congenital anomalies for fractures if the image contains linear bony lucency extending into the cortex.

FIG 10. Fracture-negative, radiologist true-negative, CNN false-positive case example. Axial cervical spine CT image demonstrates a postsurgical defect involving the right lamina secondary to laminoplasty. *Red arrow* demarcates postsurgical defect. This case example illustrates a drawback of the algorithm to fail to differentiate postsurgical changes from fractures.

contained a short interval follow-up cervical spine MR imaging. This decision was made to ensure the veracity of our ground truth analysis because the group of interpreting radiologists in our study was diverse and had individuals with various experience in cervical spine trauma evaluation.

The CNN in our study demonstrated an accuracy of 92% compared with 96% for the radiologists, underscoring the capability of the CNN at fracture detection. In addition, time from image acquisition to CNN analysis was considerably shorter than the time from image acquisition to radiologist report finalization emphasizing the value of the CNN in worklist prioritization. This benefit would be of greater value to high-volume practices that may have even longer radiology interpretation times. There is tremendous potential for worklist prioritization to improve patient outcomes by decreasing time to diagnosis and therapeutic intervention for unstable fractures.

The sensitivity of the CNN (79%) is lower than that of the radiologists (93%). CNN output should therefore be appraised after the radiologist's imaging review. Further work to improve CNN sensitivity is particularly important if CNNs are to become widely accepted as valuable worklist prioritization tools. Importantly, the clinically more useful parameters of PPV and NPV were comparable between the CNN and radiologists in our dataset consisting of a high fracture prevalence.

Our review of the few CNN false-negative examinations demonstrates that the locations of CNN misses closely match those of radiologists. Knowledge of this is important as radiologists need to be aware of the locations where the CNN performs poorly in order to subject these locations to additional scrutiny before report finalization. The few instances in which the CNN detected a fracture that the radiologist missed underscores the ability of the CNN to function as a valuable complementary tool in fracture detection that should be reviewed by the radiologist before report finalization to maximize overall fracture detection sensitivity.

Discrepant examinations reveal important limitations of the CNN. As noted in Fig 4, a severe fracture-dislocation was missed by the CNN algorithm. In addition, as noted in Fig 5, fractures characterized more by distraction rather than linear bony lucency, fractures involving the distal aspects of the spinous processes that may be mistaken for nuchal ligament calcification or ossification, and fractures located in the lower cervical spine where fine bony detail becomes poor from CT beam attenuation were also missed by the CNN. Fractures of these types must be added to the CNN training dataset as they will need to be detected if CNNs are to become increasingly valuable worklist prioritization tools.

Study design and selection bias are important limitations to our study, diminishing the generalizability of our findings. Most scans were performed at a single primary site and therefore a prospective, multicenter trial will need to be pursued next. In addition, our dataset contained a high fracture prevalence minimizing the number of clinically occult fractures and potentially falsely elevating our reported CNN and radiologist sensitivity. If

we extrapolate our sample size and values for sensitivity and specificity to a dataset that contains a much lower fracture prevalence on par with that observed in previously reported multiinstitutional cervical spine trauma trials, the PPVs for the CNN and radiologist drop below the threshold of clinical utility. Consequently, we view our results as an important first step to demonstrate CNN effectiveness in cervical spine fracture detection in a dataset with a high fracture prevalence with robust ground truth analysis, which will need to be replicated in a dataset with a lower fracture prevalence similar to routine clinical practice.

CONCLUSIONS

The CNN holds promise at both worklist prioritization and assisting radiologists in cervical spine fracture detection on CT. CNN plays an important role in prioritizing fracture-positive examinations on the worklist. Further refinements in sensitivity will improve CNN diagnostic utility. Understanding the strengths and weaknesses of the CNN is essential before its successful incorporation into clinical practice. In the evaluation of individual examinations, the current role of the CNN in fracture detection is secondary to a thorough review by a radiologist and should always be reviewed before report finalization.

Disclosures: Juan Small—*UNRELATED*: Royalties: Elsevier book royalties.

REFERENCES

1. Adams M, Chen W, Holcdorf D, et al. **Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures.** *J Med Imaging Radiat Oncol* 2019;63:27–32 CrossRef Medline
2. Urakawa T, Tanaka Y, Goto S, et al. **Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network.** *Skeletal Radiol* 2019;48:239–44 CrossRef Medline
3. Cheng CT, Ho TY, Lee TY, et al. **Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs.** *Eur Radiol* 2019;29:5469–77 CrossRef Medline
4. Chung SW, Han SS, Lee JW, et al. **Automated detection and classification of the proximal humerus fracture by using deep learning algorithm.** *Acta Orthop* 2018;89:468–73 CrossRef Medline
5. Gan K, Xu D, Lin Y, et al. **Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments.** *Acta Orthop* 2019;90:394–400 CrossRef Medline
6. Kim DH, MacKinnon T. **Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks.** *Clin Radiol* 2018;73:439–45 CrossRef Medline
7. Lindsey R, Daluiski A, Chopra S, et al. **Deep neural network improves fracture detection by clinicians.** *Proc Natl Acad Sci U S A* 2018;115:11591–96 CrossRef Medline
8. Olczak J, Fahlberg N, Maki A, et al. **Artificial intelligence for analyzing orthopedic trauma radiographs.** *Acta Orthop* 2017;88:581–86 CrossRef Medline
9. Derkatch S, Kirby C, Kimelman D, et al. **Identification of vertebral fractures by convolutional neural networks to predict nonvertebral and hip fractures: a registry-based cohort study of dual x-ray absorptiometry.** *Radiology* 2019;293:405–11 CrossRef Medline
10. Pranata YD, Wang KC, Wang JC, et al. **Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images.** *Comput Methods Programs Biomed* 2019;171:27–37 CrossRef Medline
11. Burns JE, Yao J, Summers RM. **Vertebral body compression fractures and bone density: automated detection and classification on CT images.** *Radiology* 2017;284:788–97 CrossRef Medline
12. Tomita N, Cheung YY, Hassanpour S. **Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans.** *Comput Biol Med* 2018;98:8–15 CrossRef Medline
13. Muehlematter UJ, Mannil M, Becker AS, et al. **Vertebral body insufficiency fractures: detection of vertebrae at risk on standard CT images using texture analysis and machine learning.** *Eur Radiol* 2019;29:2207–17 CrossRef Medline
14. Milby AH, Halpern CH, Guo W, et al. **Prevalence of cervical spinal injury in trauma.** *Neurosurg Focus* 2008;25:E10 CrossRef Medline
15. Minja FJ, Mehta KY, Mian AY. **Current challenges in the use of computed tomography and MR imaging in suspected cervical spine trauma.** *Neuroimaging Clin N Am* 2018;28:483–93 CrossRef Medline
16. Inaba K, Byerly S, Bush LD, et al. **Cervical spinal clearance: a prospective Western Trauma Association multi-institutional trial.** *J Trauma Acute Care Surg* 2016;81:1122–30 CrossRef Medline
17. Poonnoose PM, Ravichandran G, McClelland MR. **Missed and mismanaged injuries of the spinal cord.** *J Trauma* 2002;53:314–20 CrossRef Medline
18. Izzo R, Popolizio T, Balzano RF, et al. **Imaging of cervical spine traumas.** *Eur J Radiol* 2019;117:75–88 CrossRef Medline
19. Khanpara S, Ruiz-Pardo D, Spence SC, et al. **Incidence of cervical spine fractures on CT: a study in a large level I trauma center.** *Emerg Radiol* 2020;27:1–8 CrossRef Medline
20. Alessandrino F, Bono CM, Potter CA, et al. **Spectrum of diagnostic errors in cervical spine trauma imaging and their clinical significance.** *Emerg Radiol* 2019;26:409–16 CrossRef Medline
21. Bernstein MP, Young MG, Baxter AB. **Imaging of spine trauma.** *Radiol Clin North Am* 2019;57:767–85 CrossRef Medline