Check for updates

KJR

Assessment of a Deep Learning Algorithm for the Detection of Rib Fractures on Whole-Body Trauma Computed Tomography

Thomas Weikert, MD*, Luca Andre Noordtzij, MD*, Jens Bremerich, MD, MHBA, Bram Stieltjes, MD, PhD, Victor Parmar, MSc, Joshy Cyriac, MSc, Gregor Sommer, MD, MSc, Alexander Walter Sauter, MD

All authors: Clinic of Radiology and Nuclear Medicine, University Hospital Basel, University of Basel, Basel, Switzerland

Objective: To assess the diagnostic performance of a deep learning-based algorithm for automated detection of acute and chronic rib fractures on whole-body trauma CT.

Materials and Methods: We retrospectively identified all whole-body trauma CT scans referred from the emergency department of our hospital from January to December 2018 (n = 511). Scans were categorized as positive (n = 159) or negative (n = 352) for rib fractures according to the clinically approved written CT reports, which served as the index test. The bone kernel series (1.5-mm slice thickness) served as an input for a detection prototype algorithm trained to detect both acute and chronic rib fractures based on a deep convolutional neural network. It had previously been trained on an independent sample from eight other institutions (n = 11455).

Results: All CTs except one were successfully processed (510/511). The algorithm achieved a sensitivity of 87.4% and specificity of 91.5% on a per-examination level [per CT scan: rib fracture(s): yes/no]. There were 0.16 false-positives per examination (= 81/510). On a per-finding level, there were 587 true-positive findings (sensitivity: 65.7%) and 307 false-negatives. Furthermore, 97 true rib fractures were detected that were not mentioned in the written CT reports. A major factor associated with correct detection was displacement.

Conclusion: We found good performance of a deep learning-based prototype algorithm detecting rib fractures on trauma CT on a per-examination level at a low rate of false-positives per case. A potential area for clinical application is its use as a screening tool to avoid false-negative radiology reports.

Keywords: *Rib fractures; Trauma; Deep learning; Computer-assisted image interpretation; Computed tomography*

INTRODUCTION

Rib fractures are a common finding after thoracic trauma, occurring in approximately 40% of these patients

(1). Efforts in the polytrauma setting are focused on the detection of life-threatening conditions, such as aortic dissection and organ laceration. This fact in combination with the lack of time (2), a noisy work environment (3), satisfaction of search (4), and the frequent co-occurence of multiple traumas on whole-body computed tomography (CT) (5) lead to a significant number of missed rib fractures in this setting (6). While most rib fractures heal without surgical intervention (7), there are three reasons why it is nonetheless important to detect rib fractures: first, they are indicators of trauma-associated conditions that require immediate treatment, such as pneumothorax, and their onset can be delayed for several days (8). Second, often as a consequence of inadequate pain control, respiratory complications, such as posttraumatic pneumonia occur secondary to rib fractures (9, 10). Finally, the number and

type of rib fractures can be a basis for further treatment strategies (11, 12). Thus, an accurate detection of rib fractures on CT scans contributes to appropriate patient care (13).

To address the problem of missed rib fractures on trauma CT, some authors have proposed multiplanar (6) and rib unfolding reconstructions (14). A complementary approach comprises algorithms based on deep convolutional neural networks (DCNNs) (15) that successfully detect other findings on CT, such as myocardial infarction (16), intracranial hemorrhage (17), and acute abdominal findings (18). Given the efficiency of DCNNs in detection of findings on CT and other modalities (19-21), we hypothesized that they are also suited to detect rib fractures. While there are multiple studies on deep learning (DL)-based detection of fractures on plain radiographs (22-27), the number of algorithms detecting fractures on CT is limited. Studies on algorithms detecting vertebral body (28, 29) and skull fractures have been performed (30), but only one preliminary study dealt with the detection of rib fractures (31).

Therefore, the aim of this study was a comprehensive assessment of the diagnostic performance of a DL-based algorithm for automatic detection of rib fractures on trauma CT scans acquired within 1 year at a level-1 trauma center.

MATERIALS AND METHODS

The local ethics committee approved the study protocol and waived the requirement of obtaining informed consent (Project ID: 2019-00510).

Case Selection

We retrospectively identified all whole-body trauma CT scans and the corresponding written reports acquired at our department in 2018 with an in-house developed radiology information system/picture archiving and communication system (PACS) search engine (n = 511). Selection criteria were the procedure code and time period (January to December 2018). Examinations were classified into positive (only acute, only chronic, or both acute and chronic) and negative for rib fractures according to the written CT reports (Fig. 1). Two radiology residents (1st and 3rd year of residency) performed this classification task.

Image Acquisition

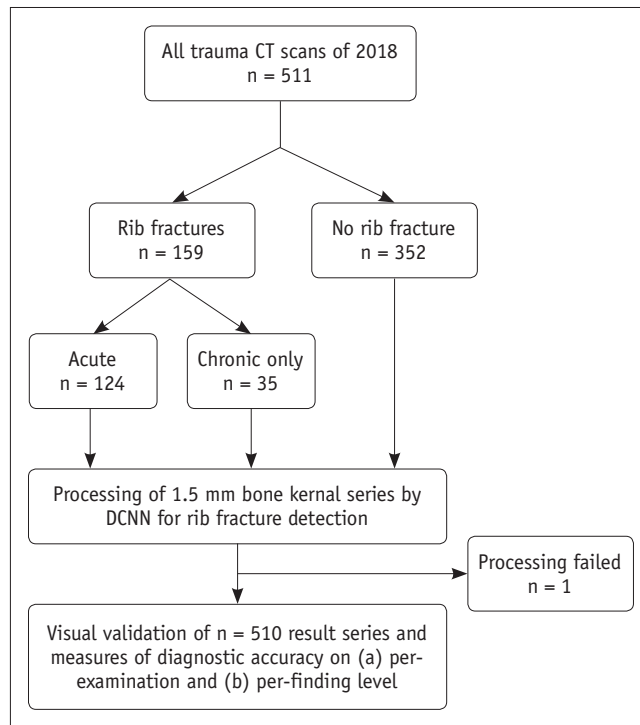Scans were acquired using three different CT scanners: Somatom Definition AS+ (n = 499; 128-slice system),

Fig. 1. Study flowchart. DCNN = deep convolutional neural network

Somatom Definition Edge (n = 5; 128-slice system), and Somatom Definition FLASH (n = 7; 2 x 128-slice system; all scanners: Siemens Healthineers, Erlangen, Germany). Scanning was performed following our standard protocol for whole-body trauma CT: patients were placed in the supine position with the scan ranging from the skull vertex to the upper thighs. Iopromide (Ultravist 370, Bayer Pharmaceuticals, Berlin, Germany) at a standard injection rate of 3.0 mL/s and a body weight-adapted volume of up to 120 mL was used as contrast agent. The peak kilovoltage was 120 kVp, and an automatic tube current modulation was performed. Transversal images in bone reconstruction kernel (70f) with a slice thickness of 1.5 mm served as the only input for the algorithm.

Index Test and Standard of Reference

The algorithm's output series with marked areas of suspected rib fractures was defined as the index test. The written CT reports established the standard of reference. These CT reports had been previously approved by a board-certified radiologist with at least 5 years of experience in emergency radiology at a level-1 trauma center. To determine the accuracy of the reports on acute fractures, we randomly selected a subset of 50 CT scans and performed a second reading without time constraints and without

knowledge of the reports.

Algorithm Characteristics

The prototype algorithm used for rib fracture detection was trained to detect acute and chronic fractures and consisted of two stages: first, a region proposal stage, which was a three-dimensional convolutional deep neural network. Its architecture was based on ResNet (Aidoc Medical, Tel Aviv, Israel). ResNets enable the training of neural networks with many layers (32). The algorithm prototype had been trained using 11455 independent chest CT scans from eight other medical institutions, acquired on 15 different scanner models. These had been reviewed by two radiologists, one making annotations and another confirming those annotations. Hyperparameter optimization included approximately 30 experiments, performed with the parallelized stochastic gradient descent using the Horovod framework (33). Experiments were processed on different servers with one, two, four, and eight NVIDIA GPUs (NVIDA, Santa Clara, CA, USA). The DCNN provided suggestions for suspected rib fractures. Subsequently, a second stage based on a Fast Region-based CNN disqualified some of the initial suggestions to reduce false-positives and selected locations for arrows indicating final findings. The output series is the original transversal series with overlaid arrows pointing at suspected rib fractures. On the internal test dataset, the performance was as follows: sensitivity 91.2% and specificity 90.7% on the per-examination level; and sensitivity 78.0% on the per-finding level.

Data Processing and Image Analysis

We performed full study data anonymization. The 1.5-mm transversal images in bone kernel (70f) were transferred to the detection algorithm. Processing the data comprised an automated cutting of the whole-body CT to the areas that displayed the ribs based on lung segmentation. These cropped series served as the only input for the core algorithm.

The output series was reviewed on a validation platform using a conventional PACS monitor (Fig. 2). A radiology resident initially reviewed all cases. Another resident and a board-certified radiologist discussed to reach a consensus on findings that were inconclusive to the first reader (e.g., does the finding display a true rib fracture or an artifact?) and all rib fractures that had been detected by the algorithm but had not been described in the written CT reports. Table 1 shows the detailed evaluation scheme for suspected findings. Acute fractures were defined as fractures without any sign of healing, such as callus formation or complete or partial consolidation of the fracture gap. A non-displaced, acute fracture was defined as a fracture with cortical disruption but maintained alignment (34). Rib fractures missed by the algorithm were marked with a bounding box.

Statistical Testing

Statistical analyses were performed with SPSS Statistics, version 22 (IBM Corp., Armonk, NY, USA) and Microsoft Excel 2010 (Microsoft Corp., Redmond, WA, USA). *P* values
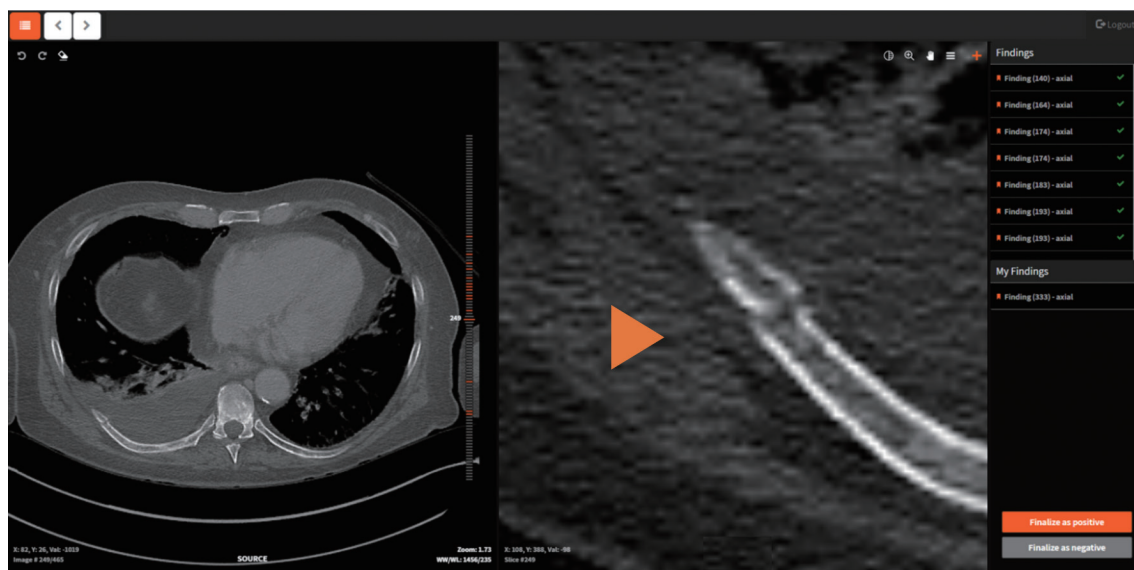

Fig. 2. Validation platform used for algorithm assessment. Original 1.5-mm transversal series of trauma CT used as input on left side. Key-image of output series with enlarged finding suspected of representing acute rib fracture marked with orange arrowhead on right side.

Table 1. Evaluation Scheme

Feature	Subfeature (If Any)	Characterization
Location	Side	Left/right
	Section	Anterior/lateral/posterior
	Number of rib	1–12
Acuteness	-	Acute/chronic
Degree of displacement	-	No displacement (= nondisplaced acute fractures + chronic fractures)/half-shaft/full-shaft/multifragmentary)
Mentioning in written report		Yes/no

less than 0.05 were considered statistically significant. We performed descriptive statistics to describe patients' age and sex. To assess if there were statistically significant differences between patients with and without fractures, we performed the Chi-squared test for sex and the Mann-Whitney U test for age. On a per-examination level, an examination was defined as true positive when the algorithm correctly identified at least one fracture in a case with rib fractures according to the report. If the algorithm did not detect any fracture in an examination with at least one fracture according to the report, this examination was classified as false negative. Cases that were classified negative for rib fractures by both the report and the algorithm were rated as true negative. We calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and F1 score for the whole dataset and subsets.

On a per-finding level ("per-fracture"), we calculated the number of false-positives per examination. To analyze the correlation of location, displacement, and acuteness with the detection rate, a binomial logistic regression with location, acuteness and degree of displacement as independent variables and detection (yes/no) as the dependent variable was performed. In this model, the following categories were used to obtain dichotomy: left/right; acute/chronic; nondisplaced/displaced (nondisplaced vs. all other categories). Exp(B) is the exponentiation of the B coefficient, which is interpreted as the odds ratio within the model (35). To further investigate the association of detection with the localization of a fracture within a rib (anterior, lateral, or posterior) and level of the fractured rib (upper = rib 1–4; middle = rib 5–8; and lower = rib 9–12), we performed Chi-squared statistics.

RESULTS

Examination Characteristics
The mean age of the patients was 58.4 ± 22.5 years.

Patients with and without rib fractures did not statistically differ in sex (χ^2 = 0.23; p = 0.63) or age (U = 19830; p = 0.44). The rate of positive examinations for rib fractures (acute and/or chronic) according to the report was 31.2%. On a per-finding level, 894 rib fractures were described in the reports. Table 2 summarizes the characteristics of the fractures. A second reading was performed for the 49 findings that had been marked as inconclusive by the first reader. Our analysis of a subset of 50 randomly selected CT scans showed that 83.3% (ten of 12) of scans showing acute rib fractures were correctly described in the corresponding reports.

Performance of the Detection Algorithm
Of 511 trauma CT scans that were performed at our department in 2018, one scan that was negative for rib fractures according to the written CT report could not be processed because of failure in automatic cropping.

Per-Examination Level
On a per-examination level, the algorithm produced 139 true-positives, 30 false-positives, 321 true-negatives, and 20 false-negatives. This corresponded to a sensitivity of 87.4% (139 of 159 scans with rib fractures according to the report detected) and specificity of 91.5% (321 of 351 scans without rib fractures correctly classified as negative) on a per-examination level. Table 3 provides more details on the performance measures. Figure 2 shows a typical example of an acute fracture of the 9th right rib that was correctly identified by the algorithm and marked with an orange arrowhead. Our sub-analysis showed that the detection sensitivity of scans that contained acute fractures was significantly higher (91.9%) compared to that of scans that contained only chronic fractures (71.4%).

Per-Finding Level
On a per-finding level, there were 587 true-positives (sensitivity: 65.7%; 95% confidence interval: 92.4–68.8)

Table 2. Characteristics of True-Positive Fractures That Were either Described in Written CT Reports (n = 894) or Additionally Detected by the Algorithm (n = 97)

Feature	Rib Fractures Described in Written CT Reports (n = 894)	Detection Rates for Subcategories in %	Rib Fractures Additionally Detected by Algorithm (n = 97)
Side			
Left	401	66.8 (268/401)	39
Right	493	64.7 (319/493)	58
Section			
Posterior	295	73.9 (218/295)	29
Lateral	383	68.7 (263/383)	41
Anterior	216	49.1 (106/216)	27
Height			
1–4	281	56.2 (158/281)	19
5–8	432	69.4 (300/432)	58
9–12	181	71.3 (129/181)	20
Acuteness			
Acute	688	67.7 (466/688)	65
Chronic	206	58.7 (121/206)	32
Degree of displacement			
No displacement	670	58.4 (391/670)	83
Half-shaft	118	88.1 (104/118)	10
Full-shaft	62	90.3 (56/62)	1
Multifragmentary	44	81.8 (36/44)	3

Table 3. Algorithm Performance on Per-Examination Level with 95% Confidence Intervals in Brackets

Sensitivity	Specificity	PPV	NPV	Accuracy	F1 Score
87.4% (81.2–92.1)	91.5% (88.0–94.2)	82.3% (76.6–86.8)	94.1% (91.4–96.0)	90.2% (87.3–92.6)	0.85

NPV = negative predictive value, PPV = positive predictive value

and 307 false-negatives. Furthermore, 97 true rib fractures (65 acute and 32 chronic) were detected by the algorithm and confirmed by consensus reading but were not mentioned in the written CT reports. The binary logistic regression model set up to ascertain the effects of laterality, displacement, and acuteness of fractures on the likelihood of detection was statistically significant (χ^2 = 69.2; p < 0.001). While laterality had no impact on detection rates within the model, displaced rib fractures were 4.84 times (Exp(B)) more likely to be detected compared to nondisplaced fractures (p < 0.001). Acute fractures were 4.60 times (Exp(B)) more likely to be detected compared to chronic fractures (p < 0.001). Furthermore, Chi-squared tests revealed significant associations between the fracture location and detection (χ^2 = 36.8, p < 0.001 for the position within a rib [anterior, lateral, posterior]; χ^2 = 16.4, p < 0.001 for the level of the rib [upper, middle, or lower]), with anteriorly and superiorly located fractures more likely to be missed. Table 2 provides detailed information on detection rates for all subcategories. The 81 false-positives translated to 0.16 false-positives per examination. Table 4 summarizes

Table 4. Morphologic Correlates of False-Positives

Anatomical Correlate	Number of Findings
Normal rib	18
Intercostal vessel	15
Breathing artifact	13
Out of bounds	11
Transition zone rib–costal cartilage	10
Fractures of other bones	10
Contrast agent artifact	3
Bone marrow calcification	1

Fractures of other bones = scapula, finger, processus transversus, Normal rib = intact rib misclassified as fracture, Out of bounds = fracture-mark with no anatomical correlation

the number and reasons for false-positives. Additionally, we found 137 "double annotations," fracture was marked multiple times by the algorithm. Figure 3 displays examples of false-positives. To further illustrate the clinical relevance of a rib fracture detection tool, Figure 4 shows an example of multiple, traumatic fractures in a 46-year-old woman after a car accident that required surgical stabilization.

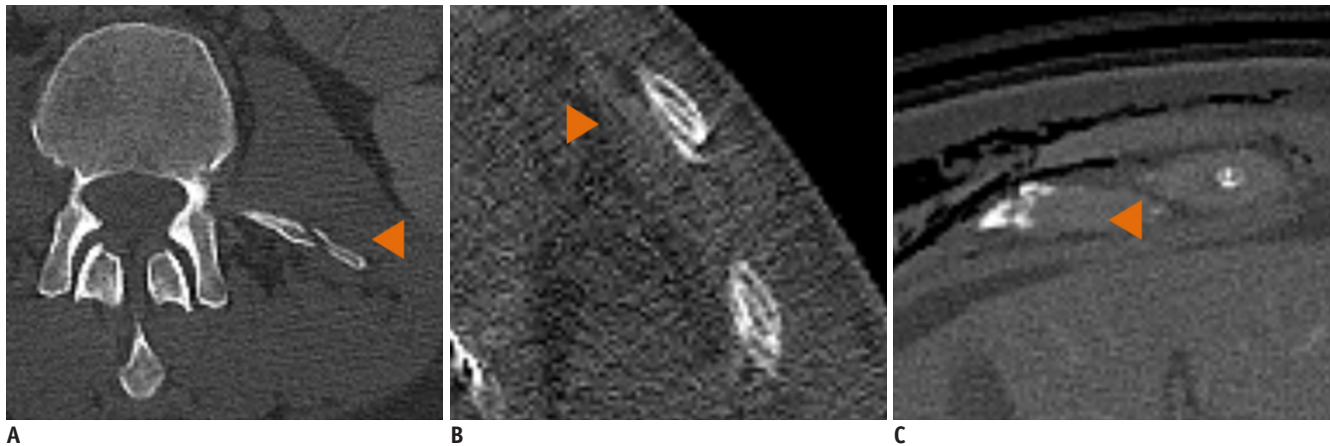A fully anonymized basic study dataset containing

Fig. 3. Three typical cases of false-positives marked with orange arrowheads, due to (A) fracture of transverse process, (B) breathing artifacts, and (C) physiological transition zone between rib and costal cartilage.
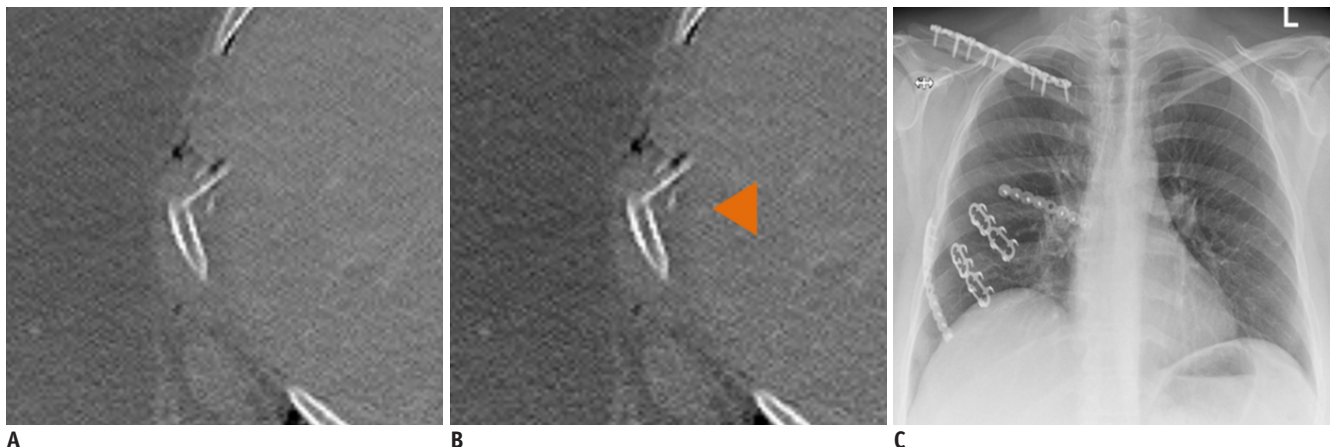

Fig. 4. Trauma CT scan of 46-year-old woman with multiple acute fractures of adjacent ribs after car accident (A) shows displaced, acute rib fracture (sixth lateral rib) that was (B) detected (orange arrowhead) and (C) subsequently required surgical stabilization.

information on the characteristics of the individual fractures and whether they were detected by the algorithm and mentioned in the radiology report can be found in Supplementary Table 1.

DISCUSSION

Our study assessed the diagnostic performance of a DL-based prototype algorithm for the automated detection of rib fractures on trauma CT. On a per-examination level, the algorithm reached a sensitivity of 87.4% and specificity of 91.5%. This is comparable to the accuracy of practicing radiologists (14). The F1 score was 0.85, and there were 0.16 false-positives per examination, mostly because of detections of intact ribs, normal intercostal vessels, and breathing artifacts. On a per-finding level, 587 of 894 fractures mentioned in the reports were detected

(sensitivity: 65.7%). Main factors associated with the correct detection of fractures by the algorithm were displacement and acuteness.

The superior performance of the algorithm on a per-examination level compared to the per-finding level may be explained by the fact that in an emergency setting, multiple rib fractures are more frequent than isolated rib fractures. In our dataset, only 9.4% of scans with rib fractures contained only one rib fracture. The detection of only part of multiple fractures is sufficient to identify a positive case correctly on the per-examination level. Due to a high NPV of 94.1%, the algorithm prototype preserves its usability as a secondary reading tool on a per-examination level.

Only one preliminary study evaluated an algorithm for the detection of rib fractures on CTs: Yan et al. (31), as well deployed CNN, yielding a sensitivity of 95.0% and a significantly lower PPV of 55.7% for the detection of rib

fractures, tested on a set of 244 fractures. The measure of false-positives and false-negatives per case was not reported.

Our results are comparable to those of other researchers investigating the performance of algorithms for the detection of fractures of other bones on CT. In their study on the performance of a support vector machine for detection of vertebral body fractures, Burns et al. (28) found a sensitivity of 81.3% on a per-finding level and 2.7 false-positives per case. The number of false-positives per case that we found was much smaller (0.16), which translates to a better usability in clinical workflows. Bar et al. (29) assessed an algorithm for detection of vertebral compression fractures based on a segmentation step and a patch-based CNN and reported a sensitivity of 83.9% and specificity 93.8%. While these results are similar to ours, the authors did not include information on false-positives. However, comparability is limited because each type of fracture has different characteristics and surrounding anatomical structures. We additionally found 137 rib fractures that were annotated multiple times by the algorithm. The consequences of this depend on the way the algorithm is used: if the algorithm is used to flag scans with fractures on the per-examination level, there is no consequence; if the algorithm is used to determine the exact number of fractures, and results are checked by a radiologist in a second step, the workflow efficiency is reduced; and if the detailed results are adopted without checking, this results in a wrong assessment of the number of rib fractures.

Altogether, the algorithm detected 97 acute fractures not mentioned in the written CT report. This is of interest to clinicians since Battle et al. (36) have demonstrated that a higher number of rib fractures is associated with an increased mortality, underpinning the importance of correct rib fracture detection. Moreover, we found that detection rates for fractures located anteriorly were lower than those for other locations. Interestingly, this is in line with the results of Ringl et al. (14) and might have resulted from the diagnostically challenging area of transition from the rib to the cartilage.

Our study has several limitations. First, due to the retrospective design and limited availability of fracture-specific clinical data, the results could be linked neither to clinical symptoms nor to clinical outcomes. Second, the analysis was performed on data acquired on scanners of one vendor and at one center only. Therefore, the performance

might differ across institutions and scanners. However, CT trauma protocols are highly standardized, as they are optimized to target specific clinical questions. Therefore, we do not expect a relevant bias. Third, the algorithm output was assessed by one radiologist only with a consensus reading of two radiologists in inconclusive cases. However, the fact that the assessment of patients with trauma is a task that residents learn early in their professional career supports our conclusion that it only slightly affects the validity of the study. Fourth, the reference standard was defined by the clinically approved written CT reports. We chose this definition because we consider these reports to be a valid basis for a reference standard. A complete reading of all 511 cases by multiple readers was not possible because of the substantial time required for this task. Due to the continuum of bone healing and the resulting indefinite cut-off between acute and chronic fractures, we decided to include both fracture types.

Since the algorithm provides good results on a per case level and has a high NPV, the algorithm is usable as a screening tool to flag scans with at least one suspected rib fracture. If no rib fractures were detected by the radiologist in charge during the first reading of a trauma CT scan flagged as suspicious by the algorithm, a quick second check for rib fractures might be appropriate to avoid false-negatives.

In conclusion, the algorithm we evaluated on a large dataset independent from its training data showed good diagnostic performance for the automated detection of rib fractures on whole-body trauma CT on a per-examination level. Thus, despite lower sensitivity on a per-finding level, it constitutes a foundation for a clinical decision support tool for reading assistance.

Supplementary Materials

The Data Supplement is available with this article at https://doi.org/10.3348/kjr.2019.0653.

ORCID iDs

Thomas Weikert
 https://orcid.org/0000-0001-9274-053X
Luca Andre Noordtzij
 https://orcid.org/0000-0002-2104-127X
Jens Bremerich
 https://orcid.org/0000-0002-1002-8483
Bram Stieltjes
 https://orcid.org/0000-0002-5961-802X
Victor Parmar
 https://orcid.org/0000-0003-4802-5402
Joshy Cyriac
 https://orcid.org/0000-0002-4584-0623
Gregor Sommer
 https://orcid.org/0000-0002-8952-0808
Alexander Walter Sauter
 https://orcid.org/0000-0002-6707-2258

REFERENCES

1. Sirmali M, Türüt H, Topçu S, Gülhan E, Yazici U, Kaya S, et al. A comprehensive analysis of traumatic rib fractures: morbidity, mortality and management. *Eur J Cardiothorac Surg* 2003;24:133-138
2. Sokolovskaya E, Shinde T, Ruchman RB, Kwak AJ, Lu S, Shariff YK, et al. The effect of faster reporting speed for imaging studies on the number of misses and interpretation errors: a pilot study. *J Am Coll Radiol* 2015;12:683-688
3. Park SH, Song HH, Han JH, Park JM, Lee EJ, Park SM, et al. Effect of noise on the detection of rib fractures by residents. *Invest Radiol* 1994;29:54-58
4. Berbaum KS, Franken EA, Dorfman DD, Rooholamini SA, Coffman CE, Cornell SH, et al. Time course of satisfaction of search. *Invest Radiol* 1991;26:640-648
5. Banaste N, Caurier B, Bratan F, Bergerot JF, Thomson V, Millet I. Whole-body CT in patients with multiple traumas: factors leading to missed injury. *Radiology* 2018;289:374-383
6. Cho SH, Sung YM, Kim MS. Missed rib fractures on evaluation of initial chest CT for trauma patients: pattern analysis and diagnostic value of coronal multiplanar reconstruction images with multidetector row CT. *Br J Radiol* 2012;85:e845-e850
7. Mayberry JC, Schipper PH. *Traumatic rib fracture: conservative therapy or surgical fixation?*. In: Ferguson M, ed. *Difficult decisions in thoracic surgery*. London: Springer, 2011:489-493
8. Lu MS, Huang YK, Liu YH, Liu HP, Kao CL. Delayed pneumothorax complicating minor rib fracture after chest trauma. *Am J Emerg Med* 2008;26:551-554
9. Ho SW, Teng YH, Yang SF, Yeh HW, Wang YH, Chou MC, et al. Risk of pneumonia in patients with isolated minor rib fractures: a nationwide cohort study. *BMJ Open* 2017;7:e013029
10. Tanaka H, Yukioka T, Yamaguti Y, Shimizu S, Goto H, Matsuda H, et al. Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients. *J Trauma* 2002;52:727-732; discussion 732
11. Bemelman M, de Kruijf MW, van Baal M, Leenen L. Rib fractures: to fix or not to fix? An evidence-based algorithm. Korean *J Thorac Cardiovasc Surg* 2017;50:229-234
12. de Jong MB, Kokke MC, Hietbrink F, Leenen LPH. Surgical management of rib fractures: strategies and literature review. *Scand J Surg* 2014;103:120-125
13. Murphy CE , Raja AS, Baumann BM, Medak AJ, Langdorf MI, Nishijima DK, et al. Rib fracture diagnosis in the Panscan era. *Ann Emerg Med* 2017;70:904-909
14. Ringl H, Lazar M, Töpker M, Woitek R, Prosch H, Asenbaum U, et al. The ribs unfolded-a CT visualization algorithm for fast detection of rib fractures: effect on sensitivity and specificity in trauma patients. *Eur Radiol* 2015;25:1865-1874
15. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. Deep learning in medical imaging: general overview. *Korean J Radiol* 2017;18:570-584
16. Mannil M, von Spiczak J, Manka R, Alkadhi H. Texture analysis and machine learning for detecting myocardial infarction in noncontrast low-dose computed tomography: unveiling the invisible. *Invest Radiol* 2018;53:338-343
17. Prevedello LM, Erdal BS, Ryu JL, Little KJ, Demirer M, Qian S, et al. Automated critical test findings identification and online notification system using artificial intelligence in imaging. *Radiology* 2017;285:923-931
18. Winkel DJ, Heye T, Weikert TJ, Boll DT, Stieltjes B. Evaluation of an AI-based detection software for acute findings in abdominal computed tomography scans: toward an automated work list prioritization of routine CT examinations. *Invest Radiol* 2019;54:55-59
19. Alkadi R, Taher F, El-baz A, Werghi N. A deep learning-based approach for the detection and localization of prostate cancer in T2 magnetic resonance images. *J Digital Imaging* 2019;32:793-807
20. Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal* 2017;35:303-312
21. Cicero M, Bilbily A, Colak E, Dowdell T, Gray B, Perampaladas K, et al. Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs. *Invest Radiol* 2017;52:281-287
22. Yahalomi E, Chernofsky M, Werman M. *Detection of distal radius fractures trained by a small set of X-ray images and faster R-CNN*. In: Arai K, Bhatia R, Kapoor S, eds. *Intelligent computing*. Cham: Springer, 2019:971-981
23. Thian YL, Li Y, Jagmohan P, Sia D, Chan VEY, Tan RT.

Convolutional neural networks for automated fracture detection and localization on wrist radiographs. *Radiol Artif Intell* 2019;1:e180001

24. Kim DH, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clin Radiol* 2018;73:439-445

25. Starosolski ZA, Kan H, Annapragada AV. CNN-based radiographic acute tibial fracture detection in the setting of open growth plates. bioRxiv, 2019. Available at: https://doi.org/10.1101/506154. Accessed August 25, 2019

26. Kitamura G, Chung CY, Moore BE. Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation. *J Digit Imaging* 2019;32:672-677

27. Lindsey R, Daluiski A, Chopra S, Lachapelle A, Mozer M, Sicular S, et al. Deep neural network improves fracture detection by clinicians. *Proc Natl Acad Sci U S A* 2018;115:11591-11596

28. Burns JE, Yao J, Muñoz H, Summers RM. Automated detection, localization, and classification of traumatic vertebral body fractures in the thoracic and lumbar spine at CT. *Radiology* 2016;278:64-73

29. Bar A, Wolf L, Amitai OB, Toledano E, Elnekave E. Compression fractures detection on CT. *Medical Imaging 2017: Computer-Aided Diagnosis* 2017;10134:1013440

30. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Development and validation of deep learning algorithms for detection of critical findings in head CT scans [updated April 2018]. Cornell University, 2018. Available at: https://arxiv.org/abs/1803.05854. Accessed August 25, 2019

31. Yan L, Chuan X, Xia C, Wang S, Chen K. *Deep learning for automatic detection of fractures on chest CT scans after blunt trauma (number: B-0566)*. ECR 2019 (European Congress of Radiology);2019 February 27-March 3;Vienna, Austria

32. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Cornell University, 2015. Available at: https://arxiv.org/abs/1512.03385. Accessed August 25, 2019

33. Sergeev A, Del Balso M. Horovod: fast and easy distributed deep learning in TensorFlow [updated February 2018]. Cornell University, 2018. Available at: https://arxiv.org/abs/1802.05799. Accessed August 25, 2019

34. Talbot BS, Gange CP, Chaturvedi A, Klionsky N, Hobbs SK, Chaturvedi A. Traumatic rib injury: patterns, imaging pitfalls, complications, and treatment. *Radiographics* 2017;37:628-651

35. Park HA. An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. *J Korean Acad Nurs* 2013;43:154-164

36. Battle CE, Hutchings H, Evans PA. Risk factors that predict mortality in patients with blunt chest wall trauma: a systematic review and meta-analysis. *Injury* 2012;43:8-17