Title
The RSNA Cervical Spine Fracture CT Dataset.

Journal
Radiology: Artificial Intelligence, 5(5)

Authors
Ball, Robyn
Gumeler, Ekim
Yeom, Kristen
et al.

Publication Date
2023-09-01

DOI
10.1148/ryai.230034

Peer reviewed

The RSNA Cervical Spine Fracture CT Dataset

Hui Ming Lin, HBSc • Errol Colak, MD • Tyler Richards, MD • Felipe C. Kitamura, MD, PhD • Luciano M. Prevedello, MD, MPH • Jason Talbott, MD, PhD • Robyn L. Ball, PhD • Ekim Gumeler, MD • Kristen W. Yeom, MD • Mohammad Hamghalam, PhD • Amber L. Simpson, PhD • Jasna Strika, MD • Deniz Bulja, MD • Salita Angkurawaranon, MD, PhD • Almudena Pérez-Lara, MD, PhD • María Isabel Gómez-Alonso, MD • Johanna Ortiz Jiménez, MD • Jacob J. Peoples, PhD • Meng Law, MD, MBBS • Hakan Dogan, MD • Emre Altinmakas, MD • Ayda Youssef, MD • Yasser Mahfouz, MD, FRCR • Jayashree Kalpathy-Cramer, PhD • Adam E. Flanders, MD • for the RSNA-ASSR-ASNR Annotators and the Dataset Curation Contributors[1]

From the Department of Medical Imaging, St Michael's Hospital, Unity Health Toronto, 30 Bond St, Toronto, ON, Canada M5B 1W8 (H.M.L., E.C.); Department of Medical Imaging, University of Toronto, Toronto, Ontario, Canada (E.C.); Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, Utah (T.R.); Dasa, Universidade Federal de São Paulo (Unifesp), São Paulo, Brazil (F.C.K.); Department of Radiology, The Ohio State University, Columbus, Ohio (L.M.P.); Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, Calif (J.T.); The Jackson Laboratory, Bar Harbor, Maine (R.L.B.); Department of Radiology, Hacettepe University, Ankara, Turkey (E.G.); Standard School of Medicine, Stanford University, Stanford, Calif (K.W.Y.); School of Computing (M.H., A.L.S., J.J.P.), Department of Biomedical and Molecular Sciences (A.L.S.), and Department of Diagnostic Radiology (J.O.J.), Queen's University, Kingston, Ontario, Canada; Department of Biomedical Engineering, Qazvin Branch, Islamic Azad University, Qazin, Iran (M.H.); Department of Radiology, Cantonal Hospital Zenica, Zenica, Bosnia and Herzegovina (J.S.); Clinic of Radiology, Clinical Center University of Sarajevo, Sarajevo, Bosnia and Herzegovina (D.B.); Department of Radiology, Chiang Mai University, Chiang Mai, Thailand (S.A.); Department of Radiology, Hospital Regional Universitario de Málaga, Málaga, Spain (A.P.L.); Department of Radiology, Hospital Quirónsalud Málaga, Málaga, Spain (M.I.G.A.); Department of Radiology and Nuclear Medicine, Alfred Health, Monash University, Melbourne, Australia (M.L.); Department of Radiology, Koç University School of Medicine, Istanbul, Turkey (H.D., E.A.); Department of Diagnostic, Molecular and Interventional Radiology, Icahn School of Medicine at Mount Sinai, New York, NY (E.A.); Department of Radiology, National Cancer Institute, Cairo University, Cairo, Egypt (A.Y.); Department of Radiology, Sultan Qaboos University Hospital, Muscat, Oman (Y.M.); Department of Ophthalmology, University of Colorado Anschutz Medical Campus, Aurora, Colo (J.K.C.); Department of Radiology and Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Mass (J.K.C.); and Department of Radiology, Division of Neuroradiology, Thomas Jefferson University, Philadelphia, Pa (A.E.F.). Received February 6, 2023; revision requested April 11; revision received July 17; accepted August 10.
Address correspondence to H.M.L. (email: *Hui-Ming.Lin@unityhealth.to*).

E.C. supported by the Odette Professorship in Artificial Intelligence for Medical Imaging, St Michael's Hospital, Unity Health Toronto. M.H., A.L.S., and J.J.P. supported in part by the National Institutes of Health (grant R01 CA233888).

[1]The RSNA-ASSR-ASNR Annotators and the Dataset Curation Contributors are listed at the end of this article.

Conflicts of interest are listed at the end of this article.

Supplemental material is available for this article.

Radiology: Artificial Intelligence 2023; 5(5):e230034 • https://doi.org/10.1148/ryai.230034 • Content codes: AI CT NR •

Cervical spine injuries are a common form of traumatic injury, affecting more than 3 million patients per year in North America (1). Cervical fractures can lead to substantial disability, as 10%–11% of all cervical spinal fractures result in spinal cord injury (2). In the United States alone, more than 1 million patients are evaluated for suspected cervical spine injury annually (3). Most of these cases are assessed in the emergent setting and, in the adult population, almost exclusively by using CT, because of its inherent improved quality and coverage relative to plain film radiography.

Interpretation of cervical spine CT can prove challenging, especially in the older population, as images are commonly confounded by superimposed degenerative disease and osteoporosis, making fracture detection more complex. Given the relatively high incidence of cervical injury in trauma patients and the potential for high morbidity, there is a need for fast and accurate diagnosis. This provides an excellent clinical use case for the assistance of an artificial intelligence (AI) algorithm. Although a few cervical spine fracture algorithms have been developed (4–6), most have limited geographic representation within the training data, restricting model generalizability. Even a model trained on a multi-institution dataset (6) had very limited diagnostic accuracy when used in practice on external datasets (7). Additionally, the lack of publicly available, expertly annotated cervical spine fracture datasets hinders further improvements in model performance using recently developed machine learning algorithms.

The Radiological Society of North America (RSNA) collaborated with the American Society of Neuroradiology (ie, ASNR) and the American Society of Spine Radiology (ie, ASSR) to create the largest publicly available, multi-institutional and multinational expert-labeled dataset of cervical spine fracture CT images for AI research, which was featured in the RSNA 2022 AI Challenge. This dataset is hosted publicly on a machine learning competition platform to help develop machine learning algorithms that can assist in the detection of cervical spine fractures. A summary of how the dataset was constructed can be found in Appendix S1.

Dataset Description and Usage

The final dataset consisted of images of the cervical spine in Digital Imaging and Communications in Medicine (DICOM) format, two comma-separated values files, and pixel-level segmentation of the cervical spine in Neuroimaging Informatics Technology Initiative (NIfTI) format. The dataset included 3112 CT scans; demographics and frequency of fractures per cervical spine level and the number of studies from each institution are shown in Table 1. The dataset is composed of 1445 studies positive for fracture (954 men, 491 women; mean age, 56.78 years ± 21.97 [SD]), of which 235 were bounding box annotated. This is supplemented with 1667 studies negative for fracture (1022 men, 645 women; mean age, 50.61 years ± 21.29). Table 2 shows the data distribution used for the Kaggle competition, with 2019 cases in the training set, 304 cases in the public test set, and 789 cases in the private test set.

Image files were organized into folders according to values stored in the Study Instance UID DICOM

Abbreviations

AI = artificial intelligence, DICOM = Digital Imaging and Communications in Medicine, NIfTI = Neuroimaging Informatics Technology Initiative, RSNA = Radiological Society of North America

Summary

This dataset is composed of cervical spine CT images with annotations related to fractures; it is available at *https://www.kaggle.com/competitions/rsna-2022-cervical-spine-fracture-detection/*.

Key Points

- This is, to our knowledge, the largest publicly available adult cervical spine fracture CT dataset, with contributions from 12 institutions across nine countries and six continents.
- This dataset includes medical images, segmentations, and expert annotations from a large cohort of radiologists with subspecialist expertise in spine imaging.
- This dataset was used successfully for the Radiological Society of North America 2022 Cervical Spine Fracture Detection competition hosted on the Kaggle machine learning platform. The dataset is made freely available to the research community for noncommercial use.

Keywords

CT, Informatics, Head/Neck, Spine, Feature Detection, Diagnosis, Segmentation

attribute, a unique study-level identifier. Individual image files within each folder were named according to their position within the stack of DICOM images via the Instance Number DICOM attribute.

The train.csv file contains study-level ground truth labels for the training set. Study Instance UID was the unique study-level identifier. The patient_overall column indicated if any cervical vertebrae were fractured, while the C1–C7 columns specified each level of the cervical spine. A value of 0 indicates absence and 1 indicates presence of fracture at that level.

The train_bounding_boxes.csv file contains information regarding the fracture bounding boxes for a subset of the training set. Study Instance UID is the unique study-level identifier. The x and y columns specify the upper left-hand corner position of the bounding box, or the point closest to (0, 0). The width and height indicate the bounding box dimensions. The slice_number column indicates the image number within the stack and can be concatenated with ".dcm" to generate the DICOM file name.

The segmentation files were named according to Study Instance UID and represent a subset of the training set. The segmentation labels have values of 1 to 7 for C1 to C7 (seven cervical vertebrae), 8 to 19 for the 12 thoracic vertebrae, and 0 for everything else. All segmented studies have C1 to C7 labels with variable inclusion of thoracic labels. The provided NIfTI files consisted of segmentation in the sagittal plane, while the DICOM files were provided in the axial plane. NIfTI header information was used to determine the appropriate orientation to ensure that the DICOM image and segmentation planes matched, as demonstrated in Figure 2. Without this information, there was a risk of having the segmentations flipped in the z-axis and/or mirrored in the x-axis.

Discussion

We curated and created expert annotation of a large high-quality cervical spine fracture CT dataset from 12 institutions from six different continents, which, to our knowledge, represents the largest public dataset of cervical spine fractures currently available. Great care was also taken to ensure that data were distributed equally with respect to sex, age, contributing site, and fracture level across the training, validation, and test sets. This additional effort helped to mitigate against unexpected or untoward performance drops between training and external or internal testing. The successful production of this dataset is partially attributed to using unconventional annotation methods by means of prelabeling. Such labels were provided by contributing sites, which allowed for the redistribution of the annotation burden.

Given the size and complexity of the dataset, much time and consideration were devoted to developing an annotation strategy that maximized the use of the annotated data while avoiding overburdening our volunteer annotators. The depth of annotations from patient level to pixel level was considered. Eventually, a hybrid schema was chosen in which images from each patient in the dataset were given a study-level annotation detailing each cervical spine level that was described as fractured in the original radiologist's report or tagged as no fracture in the control dataset. A smaller subset of patient images (approximately 16% of positive fracture cases) were assigned image-level annotations, including bounding boxes enclosing all the fractured vertebral elements on a given image. This was thought to be the best strategy to optimize the effort of the volunteers to provide "just enough" useful image-level annotations in the dataset along with the large number of additional studies with patient-level annotations. Through trial and error, the most reproducible method for image-level annotations was to have the annotators draw bounding boxes first on key sections where the fracture pattern reached a relative maximum or minimum cross-sectional area. Then the annotators skipped ahead through the image stack to the next relative maximum or minimum section and interpolated the bounding boxes in between these sections.

Establishing a strong overlap between the annotators proved to be challenging. Detailed initial instruction included example bounding boxes, a document outlining the process with image examples, and an instructional video. To help ensure accurate annotation that adhered to the provided instructions, all annotators were provided practice examinations to familiarize themselves with the tools. Performance during the practice phase was evaluated based on the ground truth bounding boxes defined by the committee, and annotators were retrained as needed. The final ground truth bounding box was calculated by taking the largest sum of all individual bounding boxes (Fig 1), which focuses on the sensitivity of fracture detection. An additional subset of cases containing segmentation masks of the vertebrae was also provided so that this could be used to help train the algorithm

Table 1: Demographic Distribution and Number of Positive and Negative Studies for Fracture per Institution

	Sex					Fracture Level Distribution						
Site	M	F	Age (y)	Positive Cases	Negative Cases	C1	C2	C3	C4	C5	C6	C7
Site 1	212	115	61.37 ± 20.98 (19–97)	169	158	32	47	11	23	24	41	68
Site 2	68	67	52.74 ± 24.35 (18–101)	92	43	15	30	4	12	13	26	33
Site 3	175	112	58.57 ± 22.34 (18–97)	100	187	15	23	10	13	20	28	44
Site 4	245	108	45.86 ± 19.19 (18–92)	182	171	34	50	28	32	48	62	75
Site 5	223	133	51.03 ± 20.95 (18–95)	176	180	26	58	17	18	30	60	76
Site 6	37	29	47.83 ± 21.92 (18–92)	30	36	2	12	3	4	4	4	11
Site 7	234	188	60.22 ± 22.32 (18–104)	187	235	28	59	17	25	35	41	62
Site 8	176	60	42.91 ± 19.28 (18–92)	96	140	18	13	10	22	25	34	34
Site 9	31	8	44.1 ± 17.81 (19–72)	11	28	0	8	2	1	1	2	1
Site 10	94	55	47.38 ± 19.64 (20–90)	41	108	4	9	6	4	7	20	15
Site 11	207	80	50.83 ± 18.74 (18–93)	144	143	35	50	9	15	29	46	67
Site 12	274	181	57.17 ± 21.68 (18–99)	217	238	32	62	21	28	49	71	88
Total	1976	1136	53.78 ± 21.83 (18–104)	1445	1667	241	421	138	197	285	435	574

Note.—Age data are reported as means ± SDs, with ranges in parentheses. The frequency of fracture(s) at each cervical spine vertebra level are tallied for the positive cases. A single study may contain multiple fractures, and thus, the sum of all fracture levels may be greater than the total number of positive cases.

Table 2: Demographic and Case Distribution among Training, Public Test, and Private Test Datasets Hosted on Kaggle

	Sex					Fracture Level Distribution						
Site	M	F	Age (y)	Positive Cases	Negative Cases	C1	C2	C3	C4	C5	C6	C7
Training	1278	741	53.65 ± 21.57 (16–104)	961	1058	146	285	73	108	162	277	393
Public test	189	115	52.51 ± 20.73 (18–101)	122	182	26	32	8	7	17	30	55
Private test	509	280	53.40 ± 22.86 (18–101)	362	427	69	104	57	82	106	128	126

Note.—Age data are reported as means ± SDs, with ranges in parentheses. The frequency of fracture(s) at each cervical spine vertebra level is tallied for the positive cases.

to detect the fracture level (Fig 2). Thus, this dataset provides multimodal annotation formats of different levels: patient level, vertebra level, bounding box, and segmentation.

The decision to request an abstraction of the radiology report from contributing sites was primarily to explore a different way to reduce the cognitive effort required in the time-consuming annotation process of an entire dataset from scratch. The rationale was that the report generated at the point of care offers the most accurate assessment, as this is when the radiologist is delivering professional services and attention is most concentrated on the task at hand. Experience has shown that volunteer annotators, even under the best of circumstances, are not reviewing examinations under the same level of scrutiny as they might in the clinical environment (12). Our goal was to redistribute and "front-load" the annotation burden and use our volunteer annotators in more of a quality-control activity. This approach, in addition to the requirement of smaller batches to contribute and annotate, offered the best balance without overburdening either the data contributors or the volunteer annotators. The challenge to this method is annotator disagreement with the ground truth report. In response to this, annotators were allowed to dispute the ground truth labels, which were subsequently adjudicated by organizing committee members.

In the future, the current dataset may be optimized by increasing the number and detail of image-level annotations or possibly by adding pixel-level annotations. The value of the dataset is not limited only to cervical spine fracture detection. For example, fractures outside of the cervical spine, including skull base, upper thoracic spine, and posterior rib fractures, were all commonly encountered and could be annotated as well to enhance the value of the dataset.

While the fracture-level distribution in the dataset is imbalanced, potentially affecting algorithm training and performance, the data distribution of fractures is similar to what has been described in real-world scenarios. For example, a multicenter study evaluated blunt traumatic cervical spine fractures at 21 different institutions and found that the most frequently fractured vertebrae were C2, C6, and C7, which together accounted for 63.3% of all cervical spine fractures (13). In the RSNA 2022 Cervical Spine Fracture Detection dataset, these three levels were also the most frequently fractured (with C7 being the most common) and accounted for 62.4% of all cervical spine fractures. A real-world distribution of the data is useful for clinical implementation of a fracture detection algorithm trained on the dataset.

There are several limitations of this dataset. The strict inclusion criteria of axial noncontrast 1-mm-thick section images may limit its application to practices that have different section acquisitions or reformat their CT cervical spine scans from a postcontrast acquisition. Additionally, the dataset treats acute and chronic fractures the same; however, detection of chronic fractures may not be as clinically relevant when evaluating trauma patients. Furthermore, this dataset excluded patients who underwent prior surgery because of the challenges of streak artifacts and altered anatomy. As such, machine learning models trained using this dataset may underperform on postsurgical scans of the cervical spine. Finally, evaluation for cervical spine fractures can be challenging, especially in the setting of severe trauma, and some fractures were visualized that were not accounted for in the radiologist's report. In these cases, the radiologist's report was chosen to represent the ground truth because of the limitations of viewing these studies in retrospect on a web-based platform. This method is obviously limited compared with radiologists reading these studies in real time on high-resolution monitors within their picture archiving and communication systems environment, with clinical history and prior imaging studies available to assist in image interpretation.

In summary, the RSNA 2022 Cervical Spine Fracture Detection dataset is, to our knowledge, the largest and most geographically diverse, publicly available expert annotated dataset of cervical spine fracture CT studies. The intent of this dataset is to inspire and enable advances in machine learning research to improve the quality, efficiency, and availability of patient care worldwide. This dataset is made freely available to all researchers for noncommercial use.

Figure 1: Axial noncontrast cervical spine CT image with bounding boxes surrounding the fractured vertebrae, annotated by individual neuroradiologists (red). Ground truth bounding box (cyan) was calculated by taking the largest sum of all individual bounding boxes, representing the largest bounding box.

Acknowledgments: The authors would like to thank and acknowledge the contributions of Christopher Carr, MA, Sohier Dane, BA, Maggie Demkin, and Michelle Riopel.

Author contributions Guarantors of integrity of entire study, **H.M.L., F.C.K., M.H., J.S., D.B., M.I.G.A.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **H.M.L., E.C., T.R., F.C.K., L.M.P., K.W.Y., M.H., J.S., M.I.G.A., H.D., A.E.F.**; clinical studies, **E.C., T.R., E.G., K.W.Y., M.H., A.L.S., J.S., D.B., S.A., A.P.L., M.I.G.A., A.Y., Y.M.**; experimental studies, **H.M.L., E.C., F.C.K., L.M.P., J.T., K.W.Y., M.H., J.S., D.B., S.A., A.P.L., J.K.C.**; statistical analysis, **H.M.L., E.C., R.L.B., M.H., A.L.S., J.S.**; and manuscript editing, **H.M.L., E.C., T.R., F.C.K., L.M.P., J.T., R.L.B., E.G., K.W.Y., M.H., A.L.S., J.S., S.A., J.O.J., M.L., H.D., E.A., A.Y., J.K.C., A.E.F.**

Dataset Curation Contributors: Nitamar Abdala, Michael Brassil, Priscila Crivellaro, Allison K. Duh, Fam Ekladious, Eduardo Moreno Júdice de Mattos Farina, Mohamed Gemae, Albert Huang, Omar Islam, Nedim Kruscica, Michael Kushdilian, Robin Lee, Zamir A. Merali, Robert Moreland, Shane Natalwalla, Oleksandra Samorodova, Yekaterina Shpanskaya, Baskaran Sundaram, Suradech Suthiphosuwan, Monica Tafur, Donatella Tampieri, Jefferson Wilson, Christopher D. Witiw, Adil Zia.

RSNA-ASSR-ASNR Annotators: Gennaro D'Anna, Allison M. Grayev, Fátima Hierro, Michael D. Hollander, Ichiro Ikuta, Christie M. Lincoln, Lubdha M. Shah, Achint K. Singh, Nathan S. Doyle, Luis G. Colon Flores, Vikas Agarwal, Scott R. Andersen, Katie Bailey, Gagandeep Choudhary, Sammy Chu, Charlotte Y. Chung, Andrea S. Costacurta, Muhammad Danial, Irene Dixe de Oliveira Santo, Venkata Dola, K. Jim Hsieh, Adham Khalil, Neil U. Lall, Laurent Letourneau-Guillon, David Russell Malin, J. Ryan Mason, Mariana Sanchez Montaño, Fanny E. Moron, Jaya Nath, Xuan V. Nguyen, Jacob Ormsby, Mark C. Oswood, Ozkan Ozsarlak, Samuel Rogers, Jeffrey Rudie, Anousheh Sayah, Eric D. Schwartz, Loizos Siakallis, Neil B. Horner, Rogerio Jadjiski de Leão.

Disclosures of conflicts of interest: **H.M.L.** No relevant relationships. **E.C.** No relevant relationships. **T.R.** No relevant relationships. **F.C.K.** Consultant for MD.ai and GE HealthCare; member of the Radiological Society of North America (RSNA)

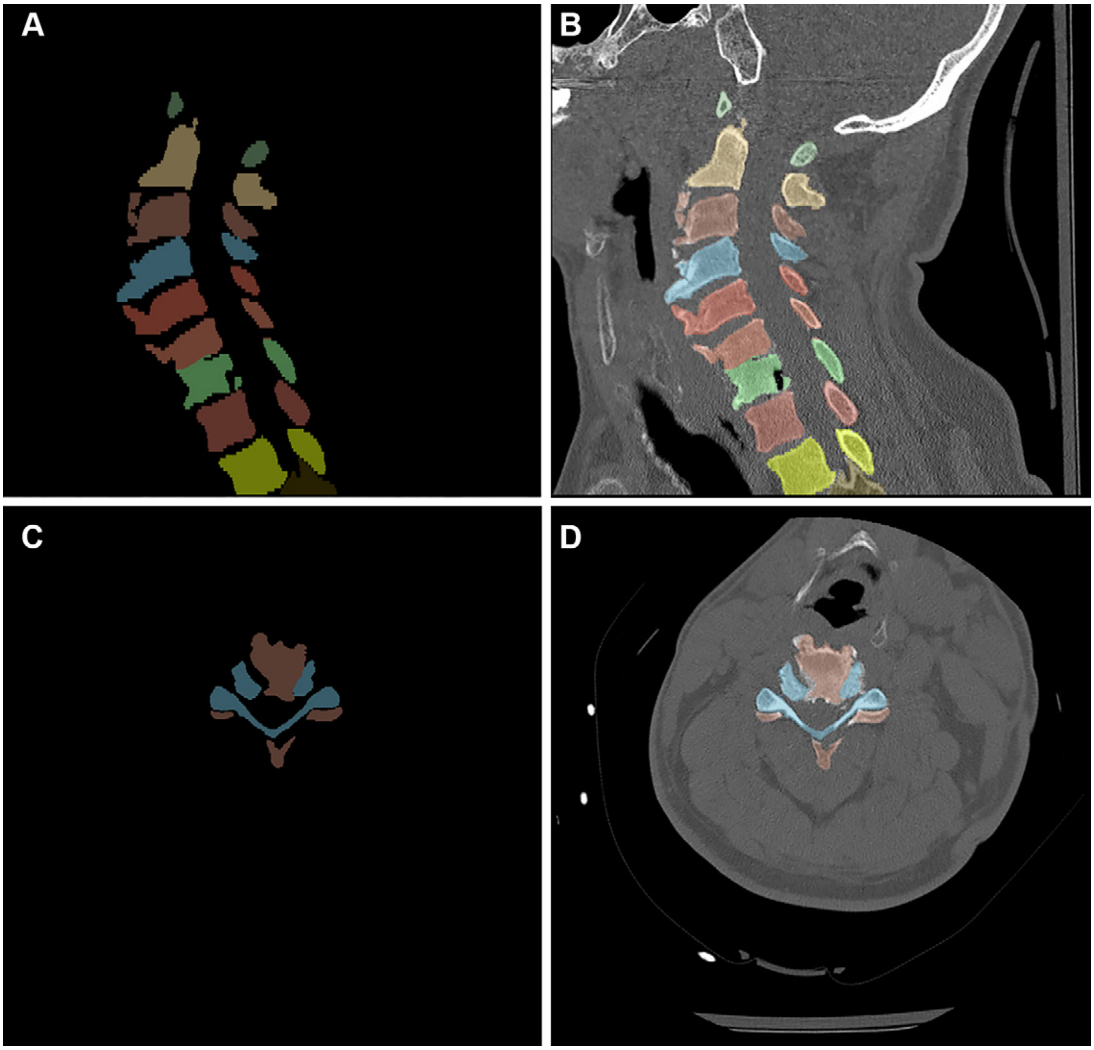

Figure 2: Example of cervical spine segmentation, with each color representing different vertebrae levels. **(A)** Image illustrates the segmentations generated in the sagittal plane. **(B)** Image depicts the segmentation mask overlaying the corresponding reconstructed sagittal DICOM section. The sagittal segmentation can be flipped onto the **(C)** axial plane, which produces **(D)** the segmentation corresponding to the original axial DICOM images. DICOM = Digital Imaging and Communications in Medicine.

Machine Learning Steering Committee member and the Society for Imaging Informatics in Medicine Machine Learning Education Subcommittee (both unpaid). **L.M.P.** Associate editor for *Radiology: Artificial Intelligence*; patents planned, issued, or pending: US-20220051060-A1, “Methods for creating privacy-protecting synthetic data leveraging a constrained generative ensemble model,” and US-20220051402-A1, “Systems for automated lesion detection and related methods.” **J.T.** Provided expert witness deposition for Phillips, Spallas, & Angstadt in October 2022, unrelated to this article. **R.L.B.** Support from RSNA to author. **E.G.** No relevant relationships. **K.W.Y.** No relevant relationships. **M.H.** No relevant relationships. **A.L.S.** Author’s lab receives funding from the National Institutes of Health (NIH); member of the advisory board for the National Cancer Institute Imaging Data Commons. **J.S.** No relevant relationships. **D.B.** No relevant relationships. **S.A.** No relevant relationships. **A.P.L.** No relevant relationships. **M.I.G.A.** No relevant relationships. **J.O.J.** No relevant relationships. **J.J.P.** Author’s lab receives funding from the NIH, which pays this author’s salary. **M.L.** No relevant relationships. **H.D.** No relevant relationships. **E.A.** No relevant relationships. **A.Y.** No relevant relationships. **Y.M.** No relevant relationships. **J.K.C.** Grants or contracts from GE HealthCare and Genentech; technology licensed to Boston AI; consulting fees from Siloam Vision; deputy editor of *Radiology: Artificial Intelligence*. **A.E.F.** Standing director, liaison for information technology, of RSNA board of directors; member of *RSNA News* editorial board.

References

1. Milby AH, Halpern CH, Guo W, Stein SC. Prevalence of cervical spinal injury in trauma. Neurosurg Focus 2008;25(5):E10.
2. Fredø HL, Bakken IJ, Lied B, Rønning P, Helseth E. Incidence of traumatic cervical spine fractures in the Norwegian population: a national registry study. Scand J Trauma Resusc Emerg Med 2014;22(1):78.
3. 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(3):483–493.
4. Dunsker SB, Zhang M, Kim L, et al. Deep-learning artificial intelligence model for automated detection of cervical spine fracture on computed tomography (CT) imaging [abstr]. J Neurosurg 2019;131(1):218.
5. Salehinejad H, Ho E, Lin HM, et al. Deep sequential learning for cervical spine fracture detection on computed tomography imaging. 2021 IEEE International Symposium on Biomedical Imaging, April 13–16, 2021.
6. Small JE, Osler P, Paul AB, Kunst M. CT cervical spine fracture detection using a convolutional neural network. AJNR Am J Neuroradiol 2021;42(7):1341–1347.
7. Voter AF, Larson ME, Garrett JW, Yu JJ. Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures. AJNR Am J Neuroradiol 2021;42(8):1550–1556.
8. Sekuboyina A, Husseini ME, Bayat A, et al. VerSe: A Vertebrae labelling and segmentation benchmark for multi-detector CT images. Med Image Anal 2021;73:102166.
9. Löffler MT, Sekuboyina A, Jacob A, et al. A vertebral segmentation dataset with fracture grading. Radiol Artif Intell 2020;2(4):e190138.
10. Liebl H, Schinz D, Sekuboyina A, et al. A computed tomography vertebral segmentation dataset with anatomical variations and multi-vendor scanner data. Sci Data 2021;8(1):284.

11. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Cham, Switzerland: Springer, 2015; 234–241.
12. Gur D, Bandos AI, Cohen CS, et al. The "laboratory" effect: comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations. Radiology 2008;249(1):47–53.
13. Goldberg W, Mueller C, Panacek E, et al. Distribution and patterns of blunt traumatic cervical spine injury. Ann Emerg Med 2001;38(1):17–21.