Radiology: Artificial Intelligence

ORIGINAL RESEARCH

SCIseg: Automatic Segmentation of Intramedullary Lesions in Spinal Cord Injury on T2-weighted MRI Scans

*Enamundram Naga Karthik, MSc** • *Jan Valošek, PhD** • *Andrew C. Smith, PT, DPT, PhD* • *Dario Pfyffer, PhD* • *Simon Schading-Sassenhausen, MD, PhD* • *Lynn Farner, MSc* • *Kenneth A. Weber II, DC, PhD* • *Patrick Freund, MD, PhD* • *Julien Cohen-Adad, PhD*

From the NeuroPoly Laboratory, Institute of Biomedical Engineering, Polytechnique Montréal, 2500 Chemin de Polytechnique, Montréal, Québec, Canada H3T 1J4 (E.N.K., J.V., J.C.A.); Mila-Quebec AI Institute, Montréal, Québec, Canada (E.N.K., J.V., J.C.A.); Department of Neurosurgery and Department of Neurology, Faculty of Medicine and Dentistry, Palacký University Olomouc, Olomouc, Czechia (J.V.); Department of Physical Medicine and Rehabilitation Physical Therapy Program, University of Colorado School of Medicine, Aurora, Colo (A.C.S.); Spinal Cord Injury Center, Balgrist University Hospital, University of Zürich, Zürich, Switzerland (D.P., S.S.S., L.F., P.F.); Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, Calif (D.P., K.A.W.); Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany (P.F.); and Functional Neuroimaging Unit, CRIUGM and Centre de Recherche du CHU Sainte-Justine, Université de Montréal, Montréal, Québec, Canada (J.C.A.). Received January 3, 2024; revision requested February 29; revision received September 11; accepted October 7. **Address correspondence to** J.C.A. (email: *jcohen@polymtl.ca*).

* E.N.K. and J.V. contributed equally to this work.

Funded by the Canada Research Chair in Quantitative Magnetic Resonance Imaging (no. CRC-2020-00179), the Canadian Institutes of Health Research (no. PJT-190258), the Canada Foundation for Innovation (nos. 32454 and 34824), the Fonds de Recherche du Québec-Santé (nos. 322736 and 324636), the Natural Sciences and Engineering Research Council of Canada (no. RGPIN-2019-07244), the Canada First Research Excellence Fund (IVADO and TransMedTech), the Courtois NeuroMod project, the Quebec BioImaging Network (nos. 5886 and 35450), INSPIRED (Spinal Research, United Kingdom; Wings for Life, Austria; Craig H. Neilsen Foundation, United States), Mila-Tech Transfer Funding Program, the Association Française contre les Myopathies, the Institut pour la Recherche sur la Moelle épinière et l'Encéphale, and the National Institutes of Health Eunice Kennedy Shriver National Institute of Child Health and Development (no. R03HD094577). E.N.K. is supported by the Fonds de Recherche du Québec Nature and Technologie (FRQNT) Doctoral Training Scholarship (B2X) and in part by the FRQNT Strategic Clusters Program (no. 2020-RS4-265502)–Centre UNIQUE–Union Neurosciences & Artificial Intelligence–Quebec and in part by funding from the Canada First Research Excellence Fund through the TransMedTech Institute. E.N.K. also acknowledges the computational support in terms of GPU resources from Digital Research Alliance of Canada. J.V. received funding from the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant (no. 101107932) and is supported by the Ministry of Health of the Czech Republic (no. NU22-04-00024). A.C.S. is supported by the National Institutes of Health (nos. K01HD106928 and R01NS128478) and the Boettcher Foundation's Webb-Waring Biomedical Research Program. K.A.W. is supported by the National Institutes of Health (nos. K23NS104211, L30NS108301, and R01NS128478). P.F. is supported by Swiss National Science Foundation (no. 32003B_204934). D.P. is supported by Swiss National Science Foundation Postdoc.Mobility Fellowship grant (P500PM_214211) and National Institutes of Health (R01 NS109450, paid to institution).

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

Radiology: Artificial Intelligence 2025; 7(1):e240005 • https://doi.org/10.1148/ryai.240005 • Content codes: AI NR

Purpose: To develop a deep learning tool for the automatic segmentation of the spinal cord and intramedullary lesions in spinal cord injury (SCI) on T2-weighted MRI scans.

Materials and Methods: This retrospective study included MRI data acquired between July 2002 and February 2023. The data consisted of T2-weighted MRI scans acquired using different scanner manufacturers with various image resolutions (isotropic and anisotropic) and orientations (axial and sagittal). Patients had different lesion etiologies (traumatic, ischemic, and hemorrhagic) and lesion locations across the cervical, thoracic, and lumbar spine. A deep learning model, SCIseg (which is open source and accessible through the Spinal Cord Toolbox, version 6.2 and above), was trained in a three-phase process involving active learning for the automatic segmentation of intramedullary SCI lesions and the spinal cord. The segmentations from the proposed model were visually and quantitatively compared with those from three other open-source methods (PropSeg, DeepSeg, and contrast-agnostic, all part of the Spinal Cord Toolbox). The Wilcoxon signed rank test was used to compare quantitative MRI biomarkers of SCI (lesion volume, lesion length, and maximal axial damage ratio) derived from the manual reference standard lesion masks and biomarkers obtained automatically with SCIseg segmentations.

Results: The study included 191 patients with SCI (mean age, 48.1 years ± 17.9 [SD]; 142 [74%] male patients). SCIseg achieved a mean Dice score of 0.92 ± 0.07 and 0.61 ± 0.27 for spinal cord and SCI lesion segmentation, respectively. There was no evidence of a difference between lesion length (P = .42) and maximal axial damage ratio (P = .16) computed from manually annotated lesions and the lesion segmentations obtained using SCIseg.

Conclusion: SCIseg accurately segmented intramedullary lesions on a diverse dataset of T2-weighted MRI scans and automatically extracted clinically relevant lesion characteristics.

Published under a CC BY 4.0 license.

Spinal cord injury (SCI) refers to damage to the spinal cord due to traumatic or nontraumatic processes. Traumatic SCI results from acute damage to the spinal cord due to external physical factors (1,2). Most patients with traumatic SCI sustain permanent neurologic deficits such as motor and autonomic dysfunction, with devastating physical and social consequences (1). Degenerative cervical myelopathy (DCM), the most common form of nontraumatic SCI, originates from chronic mechanical compression of the spinal cord (3). Although relatively less common than traumatic lesions, ischemic SCI lesions represent up to 20% of all nontraumatic lesions (4,5) and show a similar course of recovery to traumatic SCI (6,7).

MRI scans of patients with SCI provide macrostructural information about the level of injury and intramedullary abnormalities (eg, edema and hemorrhage) and allow for the evaluation of soft tissue structures (1,3). Importantly, MRI-derived quantitative biomarkers, such as intramedullary lesion length and lesion volume, have demonstrated associations with the neurologic prognosis of patients with traumatic SCI (7–12). Particularly, smaller lesion length and area were significantly associated with better recovery of patients (7,9,10).

This copy is for personal use only. To order printed copies, contact *reprints@rsna.org*

Abbreviations

DCM = degenerative cervical myelopathy, DL = deep learning, RVE = relative volume error, SCI = spinal cord injury, SCT = Spinal Cord Toolbox, 3D = three-dimensional, 2D = two-dimensional

Summary

The proposed deep learning model accurately segmented the spinal cord and spinal cord injury lesions in a diverse, multicenter dataset of T2-weighted MRI scans.

Key Points

- The SCIseg deep learning model was developed for segmentation of the spinal cord and intramedullary lesions using a multicenter dataset of 191 patients with spinal cord injury who underwent T2-weighted MRI examinations using various scanner manufacturers and acquisition parameters.
- SCIseg achieved a mean Dice score of 0.92 ± 0.07 (SD) and 0.61 ± 0.27 for spinal cord and spinal cord injury lesion segmentation, respectively.
- There was no evidence of a difference between quantitative MRI biomarkers, namely, lesion length (P = .42) and maximal axial damage ratio (P = .16), computed from manually annotated lesions and the lesion segmentations obtained using SCIseg.

Keywords

Spinal Cord, Trauma, Segmentation, MR Imaging, Supervised Learning, Convolutional Neural Network (CNN)

Despite recent advances in the automatic processing of spinal cord MRI (13–16), robust methods for detecting quantitative MRI biomarkers in SCI are lacking. As a result, most studies involve manual identification of these biomarkers (7,11,17–20). This task is time-consuming and further hampered by interrater variability, making it impractical in clinical trials (3). Furthermore, segmentation of intramedullary SCI lesions on MRI scans poses an extremely challenging task mainly due to the evolving appearance of lesions in different injury phases (eg, acute, subacute, intermediate) (1,2). Surgical implants in postoperative MRI scans might also cause severe image artifacts. Deep learning (DL) can improve diagnosis and prognostication in SCI by automating the lesion annotation process, thereby reducing rater-specific biases and facilitating the analysis of large SCI cohorts across sites (21–23). Indeed, quantitative SCI lesion biomarkers derived from DL-based automatic segmentations have been shown to correlate well with clinical measures of motor impairment (24). Despite its numerous potential advantages, DL has not been sufficiently explored in the context of SCI (22), with no open-source methods existing to date. This gap suggests the need for an automatic biomarker identification method that takes into account the complex pathophysiology of patients with SCI, generalizes to multiple sites, and remains easily accessible by researchers.

The purpose of this study was to develop an open-source DL-based tool, SCIseg, for the automatic segmentation of the spinal cord and intramedullary lesions on T2-weighted MRI scans of patients with SCI. Model-derived quantitative MRI biomarkers of SCI, such as lesion volume, lesion length, and maximal axial damage ratio, were compared with biomarkers derived from manual reference standard. Furthermore, correlation analyses were conducted between biomarkers derived using the automatic SCIseg and clinical scores, specifically light touch, pinprick, and lower extremity motor scores.

Materials and Methods

Study Design and Patients

This retrospective study included 191 patients with SCI who underwent MRI at three sites (Balgrist University Hospital Zürich, Zürich, Switzerland [site 1]; Craig Hospital, Englewood, Colorado [site 2]; Pitié-Salpêtrière University Hospital, Paris, France [site 3]) between July 2002 and February 2023. All patients provided written informed consent following institutional review board approval and the Declaration of Helsinki. The inclusion criteria were traumatic, ischemic, or hemorrhagic SCI; presence or absence of surgical hardware; and clinical data available for analyses. Exclusion criteria were concurrent traumatic brain injury beyond concussion and significant preexisting neurologic history (ie, multiple sclerosis, transverse myelitis, cerebrovascular stroke). Patients from site 2 were clinically assessed using the international standards for the neurologic classification of SCI protocol (25) to obtain light touch, pinprick, and lower extremity motor scores, as described by Smith et al (11,26). All 191 patients from the three sites were reported previously (7,11,17,26,27). These articles used manually annotated lesion masks to study the clinical consequences of SCI and their predictive relationships with motor and sensory functions. In contrast, our study presents a DL-based tool to segment intramedullary SCI lesions automatically.

MRI Data and Reference Standard

The MRI scans were converted from Digital Imaging and Communications in Medicine (DICOM) to Neuroimaging Informatics Technology Initiative (NIfTI) format and organized according to the Brain Imaging Data Structure standard (28) at individual sites. During this curation process, all sensitive patient information was deleted. T2-weighted MRI scans with varying lesion etiologies (traumatic, ischemic, hemorrhagic), injury chronicity (subacute, intermediate, and chronic), orientations (sagittal and axial), and voxel sizes were used for this study (Fig 1, Table 1). Lesions appearing as T2-weighted signal abnormalities (hyperintense or hypointense voxels corresponding to primary contusions, secondary cytotoxic edema, or hemorrhage) were manually annotated as a single object by two raters from site 1 (D.P. and L.F.), one rater from site 2 (A.C.S.), and by two raters (E.N.K. and J.V.) for site 3 using JIM (version 7.0; Xinapse Systems) and FSLeyes (version 1.9.0; University of Oxford) image viewers. As obtaining the reference standard spinal cord segmentation masks using a fully manual approach is time-consuming, sct_deepseg_sc (29) was used to initially segment the spinal cord of patients from all three sites, followed by manual corrections wherever necessary. This semiautomatic approach was also used in previous studies (29,30).

DL Training Protocol

The model was trained in three phases (Fig 2). In the initial phase, a baseline segmentation model was trained using a labeled dataset of 78 patients with T2-weighted sagittal scans (site 1) and 64 patients with T2-weighted axial scans (site 2). We used the region-based training strategy of nnU-Net (31), in which the model learns to segment the spinal cord and the

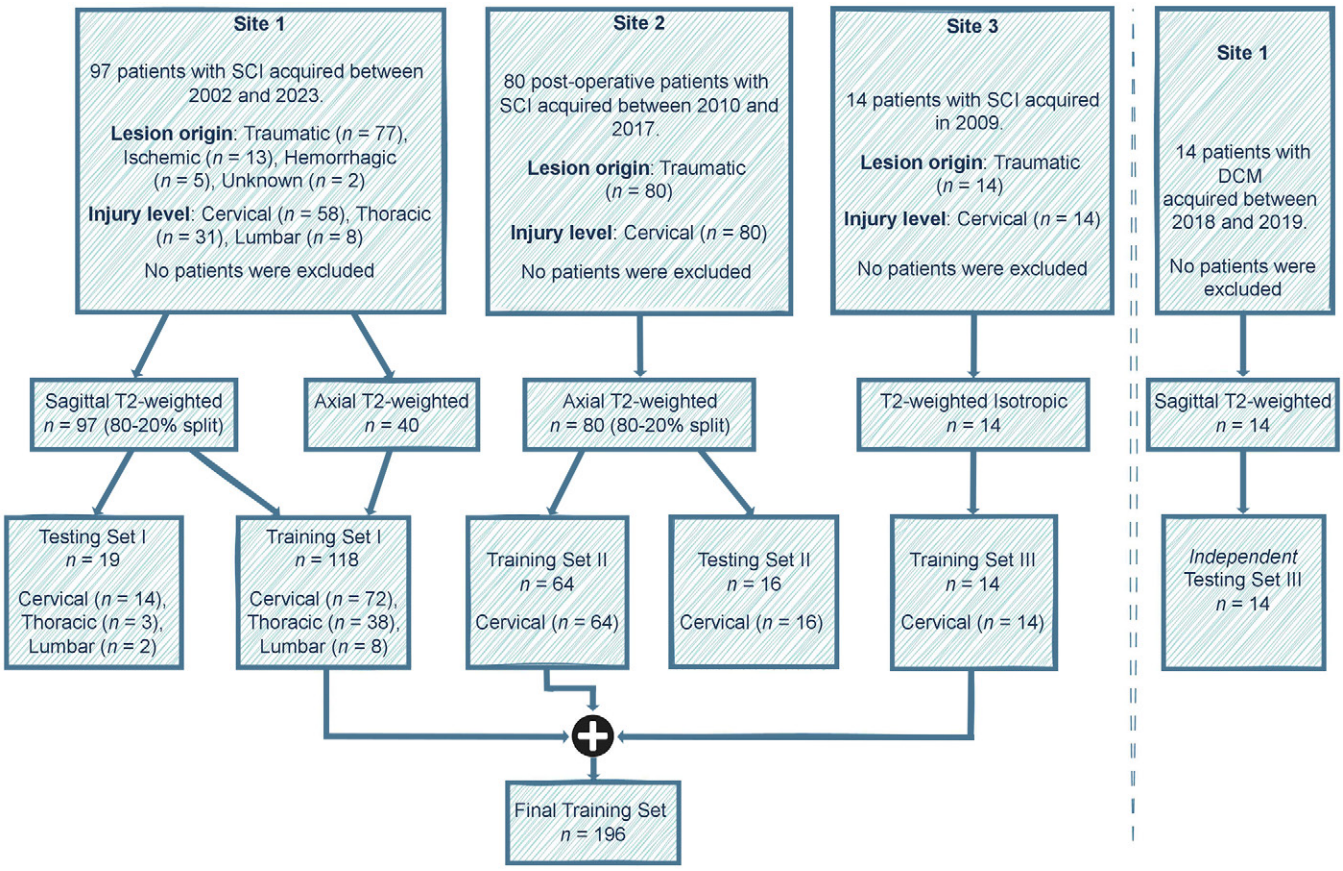

Figure 1: Study flowchart. The data included patient cohorts from three sites with heterogeneous image resolutions, orientations, and lesion etiologies. The validation set is included within the final training set. Models were evaluated independently on the test sets of site 1 and site 2, along with their evaluation on an external test set of patients with degenerative cervical myelopathy (DCM). See Table 1 for details on the MRI vendors and field strengths. SCI = spinal cord injury.

lesions simultaneously. The lesions are segmented as a single object covering hyperintense and hypointense voxels, hence containing both edema and hemorrhage. The following default nnU-Net data augmentation methods were used: random rotation, scaling, mirroring, Gaussian noise addition, Gaussian blurring, image brightness and contrast adjustment, low-resolution simulation, and gamma transformation. All scans were preprocessed in right, posterior, inferior orientation, resampled to a common resolution (0.78 × 0.56 × 0.78 mm, corresponding to the median of all image resolutions in the training set), and intensity normalized using z score normalization. The model was trained for 1000 epochs, with a batch size of two using the stochastic gradient descent optimizer with a polynomial learning rate scheduler.

For the second phase, we used the human-in-the-loop active learning strategy (32) to gradually include axial T2-weighted scans from site 1 in the training dataset. Using the phase 1 baseline model, we generated initial spinal cord and lesion predictions for unlabeled axial scans from site 1. In a subset of 40 scans, predicted segmentations underwent quality control, with two raters (E.N.K. and J.V.) manually correcting if needed. These refined segmentations were then added to the training dataset, resulting in 182 scans in the training set.

For the third training phase, to further improve our model's generalization capabilities to a wide range of image resolutions, we added a new dataset from site 3 containing 14 isotropic resolution T2-weighted sagittal scans of patients with traumatic SCI.

The final dataset consisting of 196 scans from three sites was used for training the model with the region-based strategy described above.

Evaluation Protocol

We created two independent test sets (site 1: n = 19; site 2: n = 16), following the 80/20 train/test splitting ratio (Table 1). To ensure an unbiased assessment of the model's performance and avoid overfitting, the train/test splits were done at the patient level rather than the image level, meaning that for the subset of 40 patients with sagittal and axial scans, both scans were included in the training split. For the test set, however, only the sagittal scans of each patient were used. We trained five models, each with a different train/test split using a different random seed, to avoid biasing the model toward a particular dataset split. We chose to use five random seeds instead of one fivefold cross-validation to increase the likelihood of more diverse test sets containing patients with SCI with well-defined hyperintense lesions and lesions appearing under severe metal implants. The model's performance in spinal cord segmentation was evaluated independently within each test set by comparing it with open-source methods available in the Spinal Cord Toolbox (SCT) (15): sct_propseg (33), sct_deepseg_sc (29), and the recently proposed contrast-agnostic spinal cord segmentation model (30). Due to the lack of existing state-of-the-art, open-source methods for SCI lesion segmentation, we compared the three-dimensional (3D) version of the SCIseg model with its two-dimensional (2D) version. Additionally, we tested our model on an independent cohort of 14

Table 1: Characteristics of Study Patients

Variable	Site 1 (*n* = 97)	Site 2 (*n* = 80)	Site 3 (*n* = 14)
MRI scans	137*	80	14
Sex			
Male	66	65	11
Female	25	15	2
Not reported	6	0	1
Age (y)	51.0 ± 19.1 (17–83)	45.8 ± 16.4 (15–81)	42.9 ± 16.7 (21–65)
Mean days from injury to MRI	376.2 ± 1364.4†	84.0 ± 212.1	579.1 ± 714.1
Median days from injury to MRI	41.5	21.0	407
Patients with surgical implants or hardware and their field strength	61	80	8
3 T	19	21	8
1.5 T	41	59	0
1 T	1	0	0
Lesion origin			
Traumatic	77	80	14
Ischemic	13	0	0
Hemorrhagic	5	0	0
Unknown	2	0	0
Injury level			
Cervical	58	80	14
Thoracic	31	0	0
Lumbar	8	0	0
Patients in each set			
Training set			
Cervical	44	64	14
Thoracic	28	0	0
Lumbar	6	0	0
Test set			
Cervical	14	16	0
Thoracic	3	0	0
Lumbar	2	0	0
MRI manufacturer			
Siemens	91	20	14
GE	5	60	0
Philips	1	0	0
MRI field strength			
3 T	37	21	14
1.5 T	59	59	0
1 T	1	0	0

Note.—Data are presented as numbers with percentages in parentheses, means ± SDs with ranges in parentheses, or medians. The MRI sequence parameters are as follows: site 1 sagittal: T2 weighted, voxel size 0.34 × 0.34–0.96 × 0.96 mm and section thickness, 2.2–4.8 mm; site 1 axial: T2 weighted, voxel size 0.35 × 0.35–0.78 × 0.78 mm and section thickness, 1.0–7.0 mm; site 2 axial: T2 weighted, voxel size 0.31 × 0.31–0.78 × 0.78 mm and section thickness, 3.0–6.0 mm; and site 3 isotropic: T2 weighted, voxel size 0.84 × 0.84 × 0.94–0.875 × 0.875 × 0.9 mm.

* Eight patients were followed up with more than one MRI examination.

† Five scans were acquired late after injury, resulting in a high average time for MRI examination.

patients with DCM from site 1 (unseen during training) to evaluate its generalization on patients with nontraumatic SCI (Fig 1).

Evaluation Metrics

For quantitative evaluation of model performance, we used the segmentation metrics from the open-source Anima toolkit *(https://anima.readthedocs.io/en/latest/index.html)*. For spinal cord segmentation, we calculated the Dice score and the relative volume error (RVE). For lesion segmentation, we calculated the Dice score, average surface distance, lesion-wise positive predictive value, lesion-wise sensitivity, and F1 score (34). As we trained five models on five random train/test splits (instead of fivefold cross-validation), some patients were present in more than one test set. We thus averaged the metrics across test splits for such patients.

Quantitative MRI Biomarkers

We used the Spinal Cord Toolbox's sct_analyze_lesion function to automatically compute the total lesion volume, intramedullary

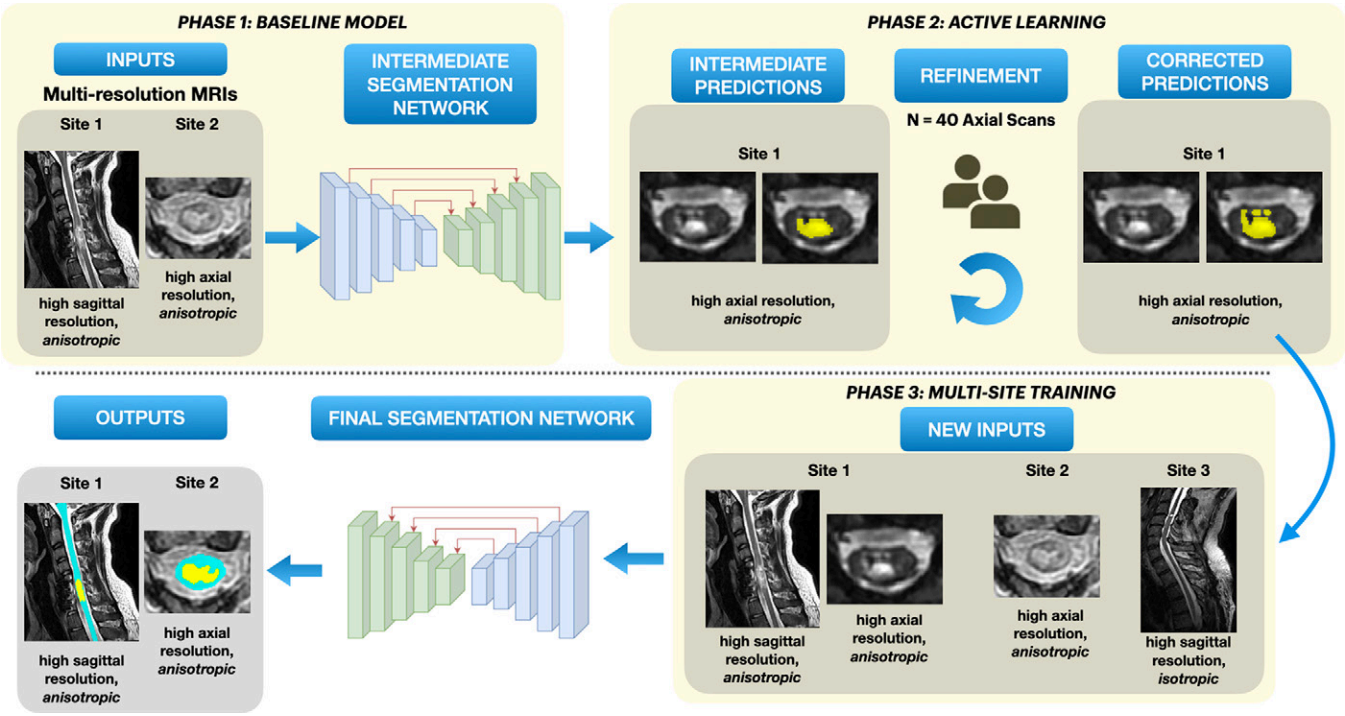

Figure 2: Overview of our segmentation approach. Phase 1: A baseline model is trained on data consisting of T2-weighted scans with axial and sagittal orientations from two sites. Phase 2: Active learning. Initial batch of automatic predictions on T2-weighted axial scans from site 2 are obtained, followed by manual corrections. Phase 3: Along with the newly corrected axial scans, isotropic T2-weighted sagittal scans from site 3 are added to the original dataset for multisite training. The final model is trained to segment both spinal cord and lesion simultaneously.

Table 2: Spinal Cord and Lesion Segmentation Performance of the Proposed SCIseg 3D Model on the Test Set

	Site 1 T2-weighted Sagittal			Site 2 T2-weighted Axial			Site 1 and Site 2
Metric	1.5 T (n = 33)	3 T (n = 46)	1.5 T and 3 T (n = 79)	1.5 T (n = 36)	3 T (n = 15)	1.5 T and 3 T (n = 51)	1.5 T and 3 T (n = 130)
Spinal cord segmentation							
Dice score*	0.88 ± 0.08	0.92 ± 0.08	0.90 ± 0.08	0.94 ± 0.04	0.94 ± 0.02	0.94 ± 0.04	0.92 ± 0.07
RVE (%)†	0.35 ± 22.32	0.17 ± 6.76	0.25 ± 15.19	1.28 ± 11.58	−1.34 ± 6.39	0.51 ± 10.33	0.35 ± 13.45
Surface distance†	0.25 ± 1.00	0.06 ± 0.18	0.14 ± 0.66	0.00 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.09 ± 0.52
Lesion segmentation							
Dice score*	0.47 ± 0.30	0.54 ± 0.30	0.51 ± 0.30	0.78 ± 0.10	0.67 ± 0.23	0.74 ± 0.15	0.61 ± 0.27
Surface distance†	5.25 ± 13.43	2.28 ± 5.43	3.51 ± 9.61	0.06 ± 0.14	0.91 ± 2.48	0.30 ± 1.34	2.17 ± 7.54
Lesion-wise PPV (%)*	51 ± 41	54 ± 45	53 ± 43	91 ± 19	81 ± 37	88 ± 26	67 ± 41
Lesion-wise sensitivity (%)*	76 ± 38	80 ± 35	79 ± 36	90 ± 20	93 ± 26	91 ± 22	84 ± 32
F1 score*	0.54 ± 0.41	0.55 ± 0.44	0.55 ± 0.43	0.88 ± 0.18	0.82 ± 0.36	0.86 ± 0.24	0.68 ± 0.39

Note.—The metrics were averaged across five random seeds and reported as means ± SDs. The best values for surface distance and relative volume error (RVE) are 0.0 and 1.0 for the rest of the metrics. PPV = positive predictive value, 3D = three-dimensional.
* The higher the value, the better.
† The lower the value, the better.

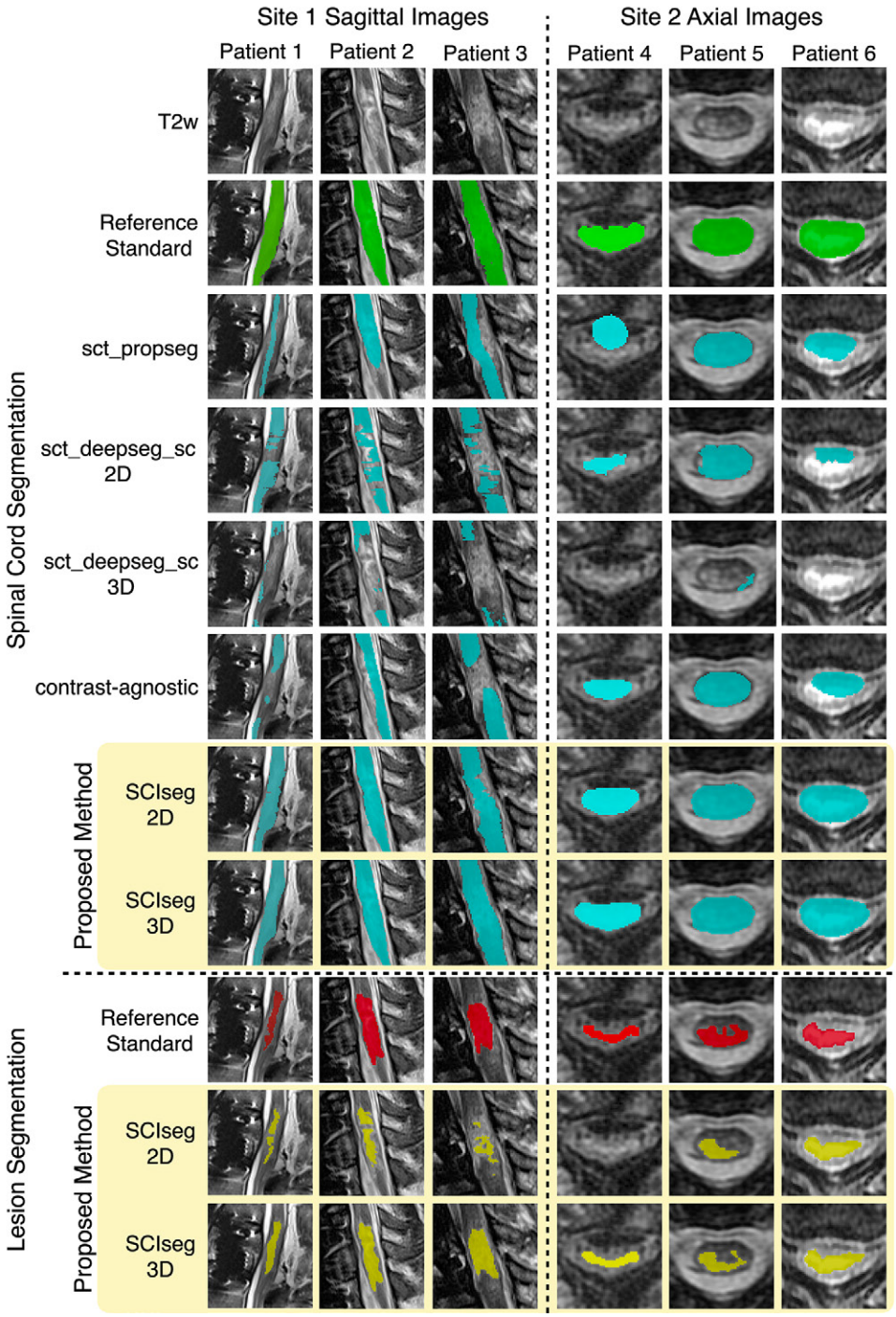

Figure 3: Comparison of SCIseg with baseline methods for the spinal cord and lesion segmentation in patients from site 1 and site 2. SCIseg 3D provides the best qualitative results for both spinal cord and lesion segmentation. Green and teal colors refer to the reference standard and automatic predictions for the spinal cord. Red and yellow refer to the reference standard and the automatic predictions for lesions. 3D = three-dimensional, 2D = two-dimensional, T2w = T2-weighted.

lesion length, and maximal axial damage ratio (11) from the manual reference standard lesion masks and the automatic predictions using the proposed SCIseg 3D model. To assess the effect of adding more training data during active learning, we computed the quantitative MRI biomarkers before (phase 1) and after active learning (phase 3). The quantitative MRI biomarkers were then averaged across five random test splits. Additionally, for site 2, we correlated the quantitative MRI biomarkers with the clinical scores (light touch, pinprick, and lower extremity motor scores).

Statistical Analysis

Statistical analysis was performed using the SciPy Python library, version 1.11.4 (35). Data normality was tested using the D'Agostino and Pearson normality test. Within-site comparisons of age and sex between patients from the testing and training sets were performed using the Mann-Whitney *U* test and the χ^2 test, respectively. Between-group comparisons (spinal cord segmentation performance SCIseg vs open-source methods; lesion segmentation performance SCIseg 2D vs SCIseg

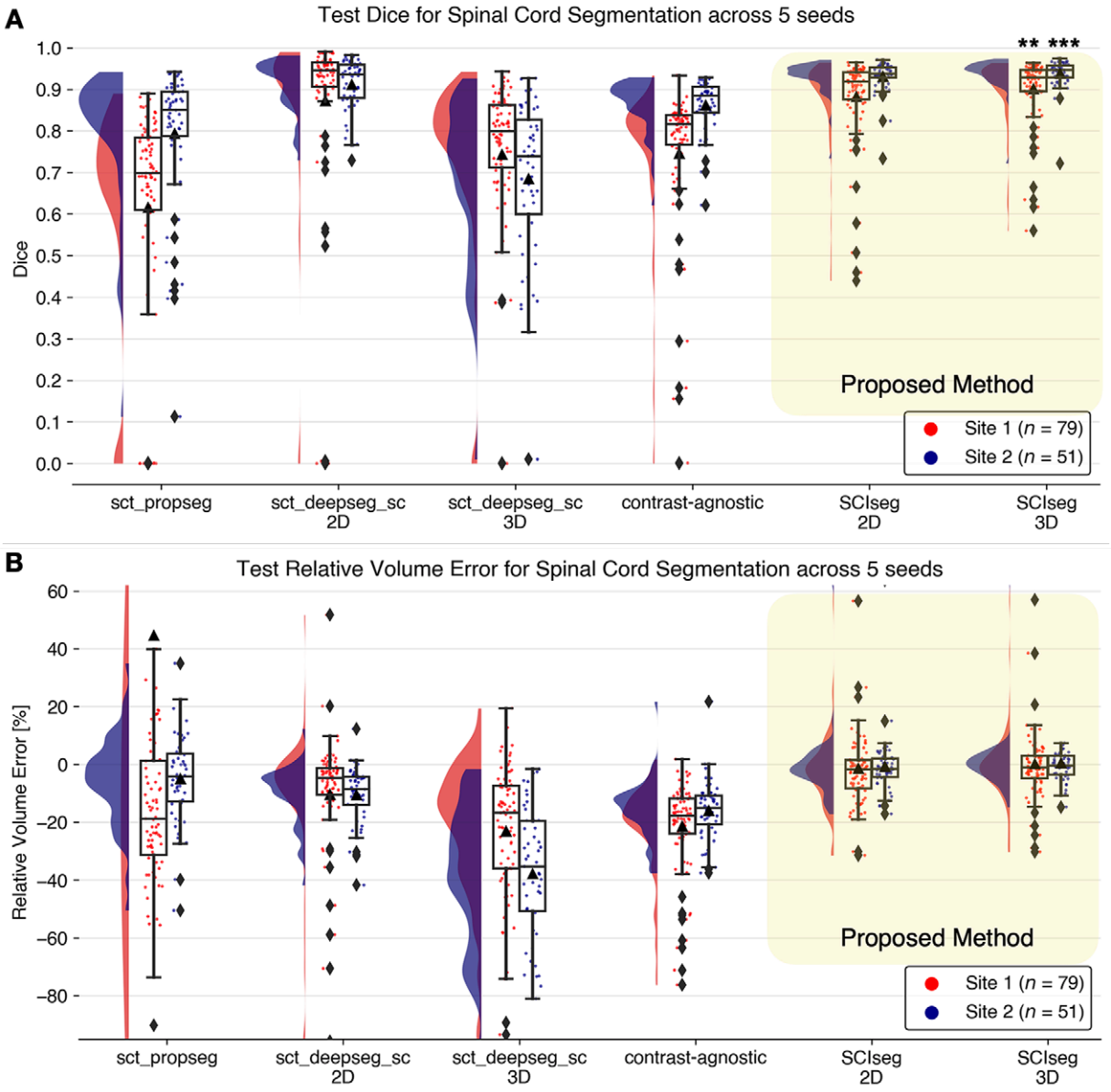

Figure 4: Raincloud plots compare the **(A)** Dice scores (best: 1; worst: 0) and **(B)** relative volume error (best: 0%) across various spinal cord segmentation methods. The numbers in the legend represent the number of test scans in each site across five different training seeds. Although the sct_deepseg_sc 2D and SCIseg 3D have similar Dice scores, the former shows a higher undersegmentation (negative relative volume error) compared with the latter. ** $P < .001$ (statistically significant for all pairs except SCIseg 3D and sct_deepseg_sc 2D). *** $P < .05$ (two-sided Bonferroni-corrected pairwise Wilcoxon signed rank test for SCIseg 3D with all baselines). 3D = three-dimensional, 2D = two-dimensional.

3D; SCIseg lesion segmentation performance before [phase 1] vs after [phase 3] active learning; manual reference standard lesion masks vs SCIseg-predicted lesions) were performed using the Wilcoxon signed rank test. The SCIseg lesion segmentation performance between sites 1 and 2 was tested using the Mann-Whitney *U* test. Correlations between clinical scores and quantitative MRI biomarkers were examined using the Spearman rank-order correlation. $P < .05$ was considered to indicate a statistically significant difference.

Code Availability

Following open science principles and to facilitate reproducibility, all codes, processing scripts, and results are open source and freely available to the SCI research community at *https://github.com/ivadomed/model_seg_sci/releases/tag/r20240716*.

Results

Patient Characteristics

A total of 191 patients (mean age, 48.1 years ± 17.9 [SD]; 142 males, 42 [22%] females, seven [4%] sex not reported) with 231 MRI scans from three sites with different lesion etiologies (traumatic, ischemic, and hemorrhagic) were included in this study (Fig 1). Ninety-seven patients were from site 1, of which 61 had surgical hardware (dorsal or ventral spondylodesis), 13 underwent decompressive surgery, and 23 did not undergo surgery. Eighty patients were from site 2, all of which had postoperative metallic stabilization. Fourteen patients were from site 3, of which eight had surgical hardware, two had decompressive surgery, and four did not undergo any surgery. Details about patient demographics,

Table 3: Comparison of Lesion Segmentation Performance between SCIseg 2D and 3D Models

	SCIseg 2D		SCIseg 3D	
Metric	Site 1 (n = 79)	Site 2 (n = 51)	Site 1 (n = 79)	Site 2 (n = 51)
Dice score*	0.36 ± 0.31	0.65 ± 0.21†	0.51 ± 0.30‡	0.74 ± 0.15†‡
Surface distance§	9.14 ± 32.43	0.54 ± 1.48†	3.51 ± 9.61‖	0.30 ± 1.34†‡
Lesion-wise PPV (%)*	33 ± 43	73 ± 35†	53 ± 43‡	88 ± 26†‖
Lesion-wise sensitivity (%)*	63 ± 46	89 ± 27#	79 ± 36‡	91 ± 22
F1 score*	0.35 ± 0.43	0.74 ± 0.33†	0.55 ± 0.43‡	0.86 ± 0.24†‖

Note.—The metrics were averaged across five different training seeds and reported as means ± SDs. The best value for surface distance is 0.0 and 1.0 for the rest of the metrics. PPV = positive predictive value, 3D = three-dimensional, 2D = two-dimensional.
* The higher the value, the better.
† Statistically significant compared with site 1. Mann-Whitney U test (P < .001).
‡ Statistically significant compared with SCIseg 2D. Wilcoxon signed rank test (P < .001).
§ The lower the value, the better.
‖ Statistically significant compared with SCIseg 2D. Wilcoxon signed rank test (P < .05).
Statistically significant compared with site 1. Mann-Whitney U test (P < .05).

injury levels, injury chronicity, and scanner types can be found in Table 1.

Eight patients from site 1 were followed up with additional MRI examinations. The final training set comprised 196 MRI scans, and the final testing set contained 35 MRI scans. There was no evidence of within-site differences in age (site 1: P = .06; site 2: P = .20) and sex (site 1: P = .71; site 2: P = .72) between patients from the testing and training sets. Patients were scanned across scanners from different manufacturers (Siemens, GE, Philips) with different field strengths (1 T, 1.5 T, 3 T). T2-weighted scans used in this study had heterogeneous image resolutions and orientations (Table 1).

Automatic Spinal Cord and Lesion Segmentation in SCI

Table 2 shows the quantitative results of SCIseg 3D on test sets of the two sites stratified by scanner strength along with their averages. As shown by the Dice scores, spinal cord segmentations from the model were stable across different data splits and magnetic field strengths despite the presence of MRI artifacts induced by spinal hardware on the scans. However, for lesion segmentation, the model performed better on site 2 compared with site 1 (mean Dice score for site 2, 0.74 ± 0.15 vs site 1, 0.51 ± 0.30), with a high SD across splits.

Comparison with Other Methods

The comparison of spinal cord and SCI lesion segmentation performance of SCIseg 3D with other methods is shown in Figures 3 and 4. The half-violin plots in Figure 4 show the distribution of the Dice scores and RVE for test scans across all seeds, and the scatterplots show the performance of the models on each test scan.

Spinal cord segmentation.— SCIseg 3D achieved the best segmentation performance (mean Dice score, 0.92 ± 0.07; mean RVE, 0.35 ± 13.45) for both sites compared with other baselines (Fig 4). For site 1, the Wilcoxon signed rank test revealed significant differences (P < .001, Bonferroni corrected) in Dice score between SCIseg 3D and all baseline methods except for sct_deepseg_sc 2D. As described in the MRI Data and Reference Standard section, this finding is the consequence of using a semiautomatic approach involving sct_deepseg_sc 2D to create the reference standard spinal cord segmentation masks. As a result, quantitative evaluations involving the reference standard masks obtained from this baseline model were inherently biased to be higher than the rest of the methods in comparison. Although there was no significant difference between SCIseg 3D and sct_deepseg_sc 2D, the SCIseg 3D model obtained spinal cord segmentations for all test cases, including those in which the baselines obtained empty predictions (shown by diamonds at Dice = 0 in Fig 4A). For site 2, statistically significant differences in Dice scores were found between SCIseg 3D and all baselines (P < .05, Bonferroni corrected). We observed more under- and oversegmented predictions by SCIseg 3D for site 1 relative to site 2 (shown by a larger spread of scatter points around RVE = 0% in Fig 4B). On visual quality control of such cases, we found that the scans contained substantial metal implants, interfering with the model's ability to fully segment the spinal cord. It must be noted that all baselines were trained specifically for segmenting the spinal cord, whereas the proposed SCIseg 3D model can segment both the spinal cord and SCI lesions simultaneously.

Lesion segmentation.— Table 3 shows a comparison between the 2D and 3D variants of the SCIseg model. The 3D model performed significantly better than the 2D model for both sites. As for the performance between sites, the 3D model's performance on site 2 was higher than that of site 1 (see Table 3 for the P values). Through visual quality control, we found that site 1 contained several patients with metal implants causing heavy image artifacts. Additionally, SCI lesions spanned different phases (acute and subacute) with various degrees of lesion hyperintensity, thus making automatic segmentation challenging.

Effect of Active Learning on Lesion Segmentation

We performed an ablation study comparing the model performance after phase 1 (training on two sites) and phase 3 (training on three sites after active learning). Figure 5A shows the

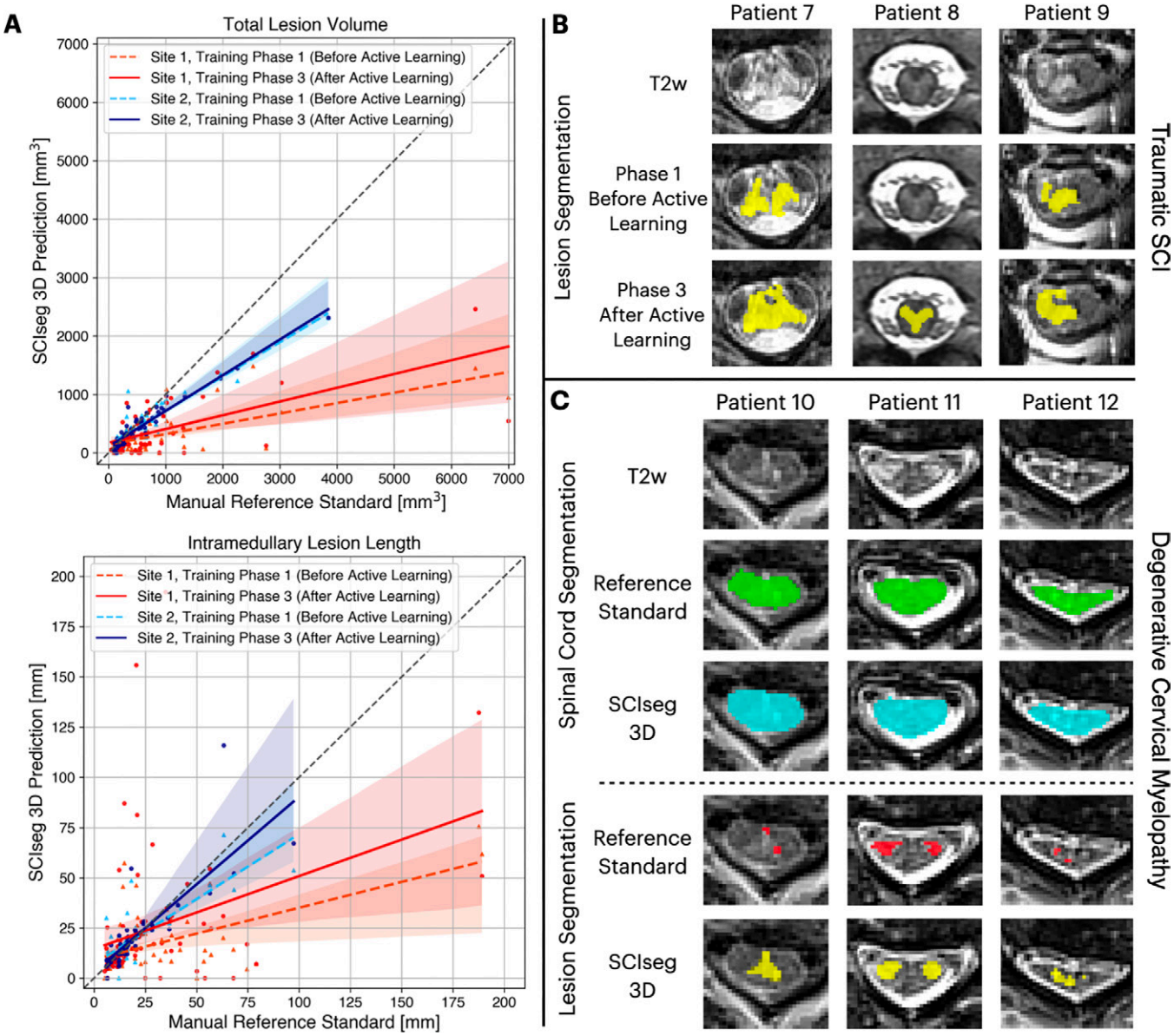

Figure 5: Comparison of model performance before and after active learning. **(A)** Correlation plots for total lesion volume (top) and intramedullary lesion length (bottom) computed from the manual reference standard lesion masks (x-axis) and lesion segmentation predictions from the proposed SCIseg 3D model (y-axis). Within each plot, colored dashed and solid lines show the agreement between the manual reference standard and automatic predictions before and after active learning, respectively, for site 1 (red and orange) and site 2 (blue and light blue). The model's predictions after active learning show a higher agreement with the manual reference standard for both sites (ie, solid lines move closer to the diagonal identity line). **(B)** SCIseg's predictions on unseen axial scans from site 2 before and after active learning. **(C)** Examples of SCIseg's generalization to patients with nontraumatic spinal cord injury (ie, degenerative cervical myelopathy). The model obtains an accurate spinal cord segmentation even at the level of severe compression (patient 12). Green and teal colors refer to the reference standard and automatic predictions for the spinal cord. Red and yellow refer to the reference standard and the automatic predictions for lesions. 3D = three-dimensional, T2w = T2-weighted.

correlation between manual reference standard and automatic predictions for total lesion volume (top) and intramedullary lesion length (bottom). For both sites, higher agreement between the manually annotated and automatically derived lesion metrics was observed for the final model after the third phase of training (ie, solid lines moving closer to the diagonal identity line). There was a statistically significant improvement after active learning (phase 3) in estimating the total lesion volume (P = .016) and the lesion length (P = .004) for site 1.

Figure 5B shows the lesion segmentation performance of our baseline model after phase 1 of training (before active learning) on unseen axial T2-weighted scans from site 1. Qualitatively, the model tends to undersegment the lesions. However, there was an overall improvement in segmentation performance when the model was trained on more data consisting of axial scans from site 1 and isotropic sagittal scans from site 3 during phase 3 of training.

Generalization to DCM

Qualitative examples of spinal cord and lesion segmentation on an independent dataset of patients with DCM whose data were unseen during model training are shown in Figure 5C. Interestingly, in cases in which the reference standard lesion masks were undersegmented, the model provided a better and more complete segmentation of the lesion. Furthermore, the spinal cord segmentations were accurate even for sections with severe spinal cord compression (Fig 5C, patient 12). Table 4 shows the Dice and

Table 4: Quantitative Evaluation of Generalizability of the SCIseg 3D Model to Patients with Nontraumatic Spinal Cord Injury

Metric	Value
Spinal cord segmentation	
Dice score*	0.95 ± 0.01
RVE (%)†	3.51 ± 4.63
Surface distance†	0.00 ± 0.00
Lesion segmentation	
Dice score*	0.46 ± 0.26
Surface distance†	1.51 ± 4.64
Lesion-wise PPV (%)*	71 ± 38
Lesion-wise sensitivity (%)*	50 ± 45
F1 score*	0.49 ± 0.44

Note.—The metrics were averaged across five random seeds and reported as means ± SDs. The best value for surface distance and relative volume error (RVE) is 0.0 and 1.0 for the rest of the metrics. PPV = positive predictive value, 3D = three-dimensional.
* The higher the value, the better.
† The lower the value, the better.

F1 scores for both spinal cord and lesion segmentations. Despite not being trained on DCM lesions, the model achieved a high mean Dice score of 0.95 ± 0.01 for spinal cord segmentation and a mean F1 score of 0.49 ± 0.44 for lesion segmentation.

Manual versus SCIseg-predicted Lesion Biomarkers

Quantitative MRI biomarkers obtained from SCIseg 3D predictions were comparable with those obtained from manually segmented lesions (Fig 6). The Wilcoxon signed rank test between SCIseg-predicted (Fig 6, green) versus manual reference standard (Fig 6, yellow) lesion biomarkers revealed a significant difference in lesion volume (P = .003) between the two groups and no evidence of a difference for lesion length (P = .42) and maximal axial damage ratio (P = .16).

Correlation between Clinical Scores and MRI Biomarkers

Figure 6 shows the correlation plots (including correlation coefficients and P values) between clinical scores and quantitative MRI biomarkers calculated from both manual reference standard lesion masks and lesions segmented using SCIseg 3D.

Discussion

This study introduced a DL-based model, SCIseg, for the automatic segmentation of the spinal cord and intramedullary lesions in patients with SCI on T2-weighted MRI scans. The model was trained and evaluated on a cohort of 191 patients with traumatic and nontraumatic SCI with 231 scans acquired using different scanner manufacturers with heterogeneous image resolutions (isotropic and anisotropic) and orientations (axial and sagittal). Patients had various lesion etiologies (traumatic, ischemic, and hemorrhagic) and lesions spread across the cervical, thoracic, and lumbar spine. SCIseg achieved a mean Dice score of 0.92 ± 0.07 for spinal cord segmentation and 0.61 ± 0.27 for SCI lesion segmentation. There was no evidence of a difference between lesion length (P = .42) and maximal axial damage ratio (P = .16) computed from manually annotated lesions and the lesion segmentations obtained using SCIseg. To the best of our knowledge, SCIseg is the first open-source, automatic method for lesion and spinal cord segmentation in SCI. It also generalizes to patients with DCM, producing accurate segmentations for both lesions and the spinal cord at the compression levels.

As the segmentation performance might be constrained by the low data quality and small dataset sizes in SCI, we showed that implementing a three-phase training strategy, including an active learning approach to progressively expand the dataset size and incorporate diverse data distributions into the training set, enhanced the model's performance. Furthermore, a region-based training strategy that jointly segments the spinal cord and the lesion was more efficient than training two individual models for spinal cord and lesion segmentation, respectively. As a result, correlation analyses between clinical scores and MRI-derived biomarkers showed statistically significant relationships for both manually annotated reference standard and automatically derived lesion masks, suggesting the SCIseg predictions can be reliably used for correlation with clinical measures in SCI.

Our cohort predominantly consisted of traumatic SCI lesions in intermediate and chronic phases, as the prevalence of ischemic and hemorrhagic lesions is typically lower (4). As chronic injuries tend to be more delineated on T2-weighted scans (2), our model learned to be sensitive to hyperintense abnormalities in the image. This sensitivity to hyperintense abnormalities also explains its ability to segment DCM lesions, which tend to be hyperintense at the site of compression. Similarly, as the injury levels in the training dataset were skewed toward the cervical spine, the model's ability to segment lumbar lesions was expected to be lower when compared with cervical or thoracic lesions. Despite the presence of metal implants causing strong image artifacts in several patients, SCIseg provided a good starting point for obtaining lesion and spinal cord segmentations instead of manual annotations from scratch.

Only a few studies exist in the literature discussing the importance of automatic segmentation in SCI scans (24,36). The study by McCoy et al (24) is most similar to ours, as it presented the first DL method for segmentation of the spinal cord and intramedullary lesions in SCI. Nevertheless, there are several important distinctions between the two studies. Although their model was trained on axial preoperative scans of patients with acute SCI from a single site, our model was trained on multisite data consisting of patients with traumatic, ischemic, and hemorrhagic SCI with different image orientations (axial and sagittal). Moreover, our model was exposed to more heterogeneous data covering different injury phases (intermediate and chronic) and therefore demonstrated better generalization to both traumatic and nontraumatic lesion etiologies. More importantly, our work is open source, further enabling reproducible, multisite studies in SCI.

Several promising avenues for future work exist. The segmentation models can be improved by using more fine-grained reference standard masks, in which the hyperintense edema and hypointense hemorrhage could be treated as separate classes. Training a model on preoperative traumatic SCI data using these reference standard masks would have a major impact on improving the initial classification of the disease and further prognostication (18). Although the model generalized reasonably well to

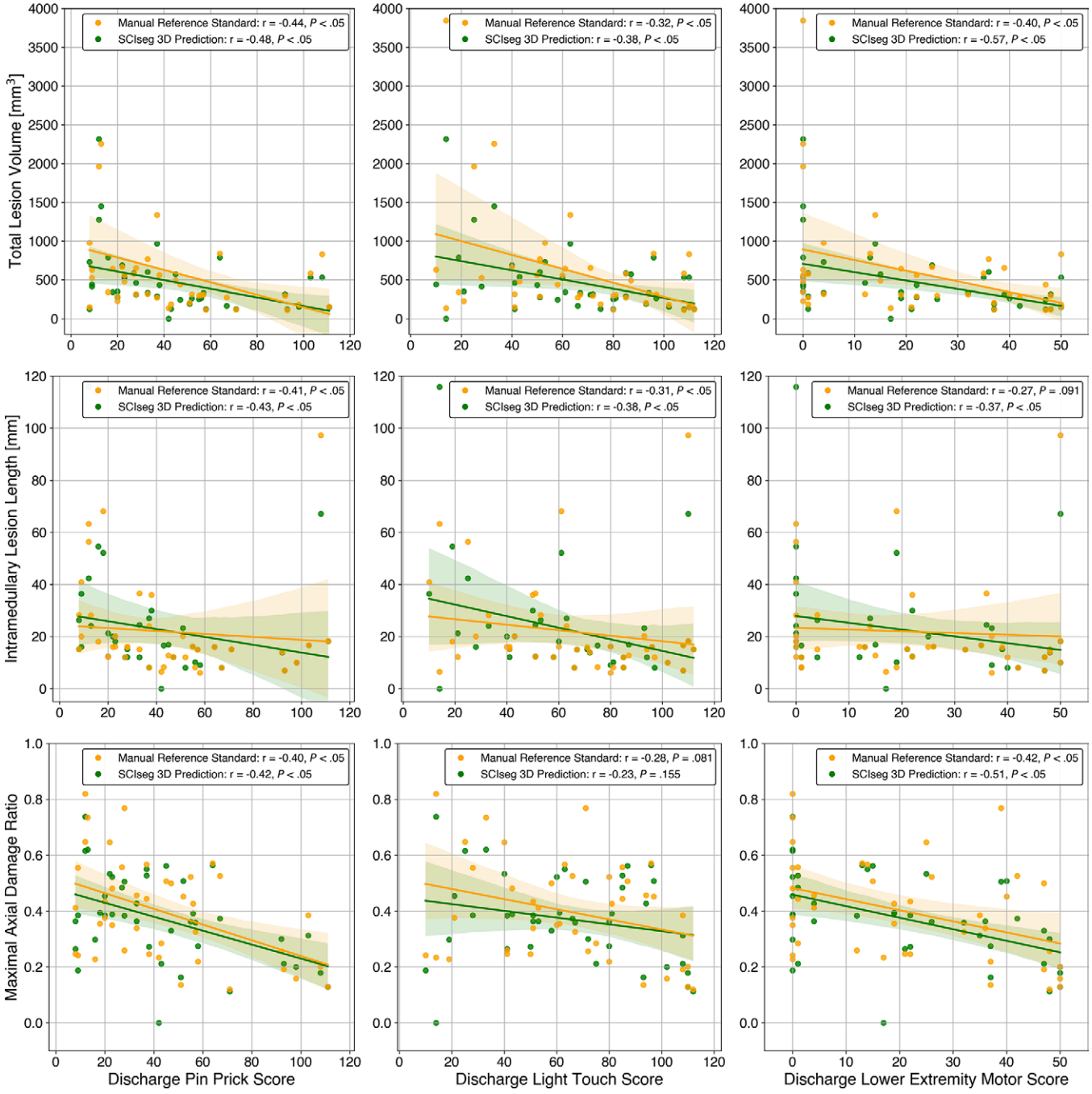

Figure 6: Correlation analysis between discharge clinical scores (x-axis) and quantitative MRI biomarkers (y-axis) for site 2. Spearman correlation coefficient and *P* value are shown in the legends of individual subplots. The Wilcoxon signed rank test between the manual reference standard lesion masks (yellow) versus automatic predictions using SCIseg 3D (green) lesion biomarkers revealed no evidence of differences for lesion length ($P = .42$) and maximal axial damage ratio ($P = .16$) but a significant difference for lesion volume ($P = .003$). 3D = three-dimensional.

DCM lesions, there is potential for improvement, particularly by adding the DCM cohort to the existing training set or by training a DL model exclusively on DCM data. Previous studies have reported the presence of hyperintense T2-weighted lesions in up to 64% of patients with DCM (13,37,38) and explored the relationship between structural and functional damages (39). Such studies would greatly benefit from an automatic DCM lesion segmentation method. Aggregating more heterogeneous data will also allow for sensitivity analysis of potential confounding factors.

This study had limitations. First, longitudinal scans from patients with follow-up examinations were treated as independent inputs for training. Although the lesion appearance evolved between sessions, resulting in nonidentical lesions (hence justifying our choice of treating them as independent inputs), the model likely could not learn the evolution of lesions across time. Second, the model's sensitivity to hyperintense abnormalities might result in false-positive segmentations in healthy controls where the spinal cord central canal is visualized. Third, our limited size of 196 MRI scans in the training set risks overfitting, given the complexity of the SCI lesion segmentation task. Although we gathered diverse data from three sites and trained five models on different train/test splits, along with extensive data augmentation to prevent potential overfitting, increasing the dataset size would further improve the model's performance and generalization. Last,

we did not analyze the interrater variability as the data were gathered from multiple sites and there were no overlapping patients across sites. Previous studies (40,41) have reported that MRI measures of spinal cord damage (eg, edema length, midsagittal tissue bridge ratio, axial damage ratio) do exhibit high to excellent levels of interrater reliability.

In conclusion, this study presented SCIseg, an automatic DL-based method, for the segmentation of the spinal cord and intramedullary lesions in SCI on T2-weighted MRI scans. The work has addressed several limitations of previous studies by including a large retrospective cohort consisting of 191 patients spanning three sites, using MRI data acquired with scanners from different manufacturers, and training a single model to segment both the spinal cord and SCI lesions. More importantly, the method has been designed to ensure reproducibility and enable large-scale, reproducible prospective studies. The model is open source and accessible via the Spinal Cord Toolbox (version 6.2 and higher). We hope that SCIseg will benefit clinicians and patients by providing additional diagnostic and prognostic information, serving as a basis for further studies assessing optimal rehabilitation from a customized patient-based perspective.

Acknowledgments: We thank Mathieu Guay-Paquet, PhD, and Joshua Newton, BEng, for their assistance with the management of the datasets and the implementation of the algorithm to the Spinal Cord Toolbox. We thank Serge Rossignol, MD, PhD, and the multidisciplinary team on locomotor rehabilitation (Regenerative Medicine and Nanomedicine, Canadian Institutes of Health Research) and all the patients. The authors thank the Digital Research Alliance of Canada for the computational resources used in this work.

Author contributions: Guarantor of integrity of entire study, **J.C.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, **E.N.K., J.V., A.C.S., J.C.A.**; clinical studies, **A.C.S., K.A.W.**; experimental studies, **E.N.K., J.V., K.A.W., J.C.A.**; statistical analysis, **E.N.K., J.V., A.C.S.**; and manuscript editing, **E.N.K., J.V., A.C.S., D.P., K.A.W., P.F., J.C.A.**

Disclosures of conflicts of interest: **E.N.K.** No relevant relationships. **J.V.** No relevant relationships. **A.C.S.** Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health – K01HD106928; Boettcher Foundation's Webb-Waring Biomedical Research Program. **D.P.** No relevant relationships. **S.S.S.** No relevant relationships. **L.F.** No relevant relationships. **K.A.W.** Grants from the National Institutes of Health paid to institution; honoraria for lecture, Norwegian Chiropractic Association; honorarium for lecture, MGH; American Academy of Pain Physicians Co-Section Editor. **P.F.** No relevant relationships. **J.C.A.** No relevant relationships.

References

1. Ahuja CS, Wilson JR, Nori S, et al. Traumatic spinal cord injury. Nat Rev Dis Primers 2017;3(1):17018.
2. Freund P, Seif M, Weiskopf N, et al. MRI in traumatic spinal cord injury: from clinical assessment to neuroimaging biomarkers. Lancet Neurol 2019;18(12):1123–1135.
3. David G, Mohammadi S, Martin AR, et al. Traumatic and nontraumatic spinal cord injury: pathological insights from neuroimaging. Nat Rev Neurol 2019;15(12):718–731.
4. Scivoletto G, Torre M, Mammone A, et al. Acute Traumatic and Ischemic Spinal Cord Injuries Have a Comparable Course of Recovery. Neurorehabil Neural Repair 2020;34(8):723–732.
5. New PW, Cripps RA, Bonne Lee B. Global maps of non-traumatic spinal cord injury epidemiology: towards a living data repository. Spinal Cord 2014;52(2):97–109. [Published correction appears in Spinal Cord 2014;52(5):417.]
6. Iseli E, Cavigelli A, Dietz V, Curt A. Prognosis and recovery in ischaemic and traumatic spinal cord injury: clinical and electrophysiological evaluation. J Neurol Neurosurg Psychiatry 1999;67(5):567–571.
7. Pfyffer D, Huber E, Sutter R, Curt A, Freund P. Tissue bridges predict recovery after traumatic and ischemic thoracic spinal cord injury. Neurology 2019;93(16):e1550–e1560.
8. Miyanji F, Furlan JC, Aarabi B, Arnold PM, Fehlings MG. Acute cervical traumatic spinal cord injury: MR imaging findings correlated with neurologic outcome--prospective study with 100 consecutive patients. Radiology 2007;243(3):820–827.
9. Dobran M, Aiudi D, Liverotti V, et al. Prognostic MRI parameters in acute traumatic cervical spinal cord injury. Eur Spine J 2023;32(5):1584–1590.
10. Huber E, Lachappelle P, Sutter R, Curt A, Freund P. Are midsagittal tissue bridges predictive of outcome after cervical spinal cord injury? Ann Neurol 2017;81(5):740–748.
11. Smith AC, Albin SR, O'Dell DR, et al. Axial MRI biomarkers of spinal cord damage to predict future walking and motor function: a retrospective study. Spinal Cord 2021;59(6):693–699.
12. Kurpad S, Martin AR, Tetreault LA, et al. Impact of Baseline Magnetic Resonance Imaging on Neurologic, Functional, and Safety Outcomes in Patients With Acute Traumatic Spinal Cord Injury. Global Spine J 2017;7(3 Suppl):151S–174S.
13. Martin AR, De Leener B, Cohen-Adad J, et al. A novel MRI biomarker of spinal cord white matter injury: T2*-weighted white matter to gray matter signal intensity ratio. AJNR Am J Neuroradiol 2017;38(6):1266–1273.
14. Badhiwala JH, Ahuja CS, Akbar MA, et al. Degenerative cervical myelopathy - update and future directions. Nat Rev Neurol 2020;16(2):108–124.
15. De Leener B, Lévy S, Dupont SM, et al. SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data. Neuroimage 2017;145(Pt A):24–43.
16. Bischof A, Papinutto N, Keshavan A, et al. Spinal Cord Atrophy Predicts Progressive Disease in Relapsing Multiple Sclerosis. Ann Neurol 2022;91(2):268–281.
17. Smith AC, Weber KA 2nd, O'Dell DR, Parrish TB, Wasielewski M, Elliott JM. Lateral Corticospinal Tract Damage Correlates With Motor Output in Incomplete Spinal Cord Injury. Arch Phys Med Rehabil 2018;99(4):660–666.
18. Mummaneni N, Burke JF, DiGiorgio AM, et al. Injury volume extracted from MRI predicts neurologic outcome in acute spinal cord injury: A prospective TRACK-SCI pilot study. J Clin Neurosci 2020;82(Pt B):231–236.
19. Vallotton K, Huber E, Sutter R, Curt A, Hupp M, Freund P. Width and neurophysiologic properties of tissue bridges predict recovery after cervical injury. Neurology 2019;92(24):e2793–e2802.
20. Pfyffer D, Vallotton K, Curt A, Freund P. Predictive Value of Midsagittal Tissue Bridges on Functional Recovery After Spinal Cord Injury. Neurorehabil Neural Repair 2021;35(1):33–43.
21. Khan O, Badhiwala JH, Wilson JRF, Jiang F, Martin AR, Fehlings MG. Predictive Modeling of Outcomes After Traumatic and Nontraumatic Spinal Cord Injury Using Machine Learning: Review of Current Progress and Future Directions. Neurospine 2019;16(4):678–685.
22. Dietz N, Vaitheesh Jaganathan, Alkin V, Mettille J, Boakye M, Drazin D. Machine learning in clinical diagnosis, prognostication, and management of acute traumatic spinal cord injury (SCI): A systematic review. J Clin Orthop Trauma 2022;35:102046.
23. Asgari Taghanaki S, Abhishek K, Cohen JP, Cohen-Adad J, Hamarneh G. Deep semantic segmentation of natural and medical images: a review. Artif Intell Rev 2021;54(1):137–178.
24. McCoy DB, Dupont SM, Gros C, et al. Convolutional Neural Network-Based Automated Segmentation of the Spinal Cord and Contusion Injury: Deep Learning Biomarker Correlates of Motor Impairment in Acute Spinal Cord Injury. AJNR Am J Neuroradiol 2019;40(4):737–744.
25. Rupp R, Biering-Sørensen F, Burns SP, et al. International Standards for Neurological Classification of Spinal Cord Injury: Revised 2019. Top Spinal Cord Inj Rehabil 2021;27(2):1–22.
26. Smith AC, O'Dell DR, Thornton WA, et al. Spinal Cord Tissue Bridges Validation Study: Predictive Relationships With Sensory Scores Following Cervical Spinal Cord Injury. Top Spinal Cord Inj Rehabil 2022;28(2):111–115.
27. Cohen-Adad J, El Mendili MM, Lehéricy S, et al. Demyelination and degeneration in the injured human spinal cord detected with diffusion and magnetization transfer MRI. Neuroimage 2011;55(3):1024–1033.
28. Gorgolewski KJ, Auer T, Calhoun VD, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 2016;3(1):160044.
29. Gros C, De Leener B, Badji A, et al. Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. Neuroimage 2019;184:901–915.
30. Bédard S, Enamundram NK, Tsagkas C, et al. Towards contrast-agnostic soft segmentation of the spinal cord. ArXiv 2310.15402 [preprint] https://arxiv.org/abs/2310.15402. Posted October 23, 2023. Accessed November 2023.
31. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18(2):203–211.

32. Budd S, Robinson EC, Kainz B. A survey on active learning and human-in-the-loop deep learning for medical image analysis. Med Image Anal 2021;71:102062.
33. De Leener B, Kadoury S, Cohen-Adad J. Robust, accurate and fast automatic segmentation of the spinal cord. Neuroimage 2014;98:528–536.
34. Commowick O, Istace A, Kain M, et al. Objective Evaluation of Multiple Sclerosis Lesion Segmentation using a Data Management and Processing Infrastructure. Sci Rep 2018;8(1):13650.
35. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020;17(3):261–272. [Published correction appears in Nat Methods 2020;17(3):352.]
36. Blanc C, Shahrampour S, Mohamed FB, de Leener B. Combining PropSeg and a convolutional neural network for automatic spinal cord segmentation in pediatric populations and patients with spinal cord injury. Int J Imaging Syst Technol 2023;33(4):1396–1405.
37. Nouri A, Martin AR, Kato S, Reihani-Kermani H, Riehm LE, Fehlings MG. The Relationship Between MRI Signal Intensity Changes, Clinical Presentation, and Surgical Outcome in Degenerative Cervical Myelopathy: Analysis of a Global Cohort. Spine 2017;42(24):1851–1858.
38. Martin AR, Tadokoro N, Tetreault L, et al. Imaging Evaluation of Degenerative Cervical Myelopathy: Current State of the Art and Future Directions. Neurosurg Clin N Am 2018;29(1):33–45.
39. Scheuren PS, David G, Kramer JLK, et al. Combined Neurophysiologic and Neuroimaging Approach to Reveal the Structure-Function Paradox in Cervical Myelopathy. Neurology 2021;97(15):e1512–e1522.
40. Smith AC, Weber KA, Parrish TB, et al. Ambulatory function in motor incomplete spinal cord injury: a magnetic resonance imaging study of spinal cord edema and lower extremity muscle morphometry. Spinal Cord 2017;55(7):672–678.
41. Cummins DP, Connor JR, Heller KA, et al. Establishing the inter-rater reliability of spinal cord damage manual measurement using magnetic resonance imaging. Spinal Cord Ser Cases 2019;5(1):20.