Radiology: Artificial Intelligence

ORIGINAL RESEARCH

# Quantitative Analysis of Neural Foramina in the Lumbar Spine: An Imaging Informatics and Machine Learning Study

*Bilwaj Gaonkar, PhD • Joel Beckett, MD, MHS • Diane Villaroman, BS • Christine Ahn, BS • Matthew Edwards, BS • Steven Moran, BS • Mark Attiah, MD, MS • Diana Babayan, BA • Christopher Ames, MD • J. Pablo Villablanca, MD • Noriko Salamon, MD • Alex Bui, PhD • Luke Macyszyn, MD, MA*

From the Departments of Neurosurgery (B.G., J.B., D.V., C. Ahn, M.E., M.A., D.B., L.M.), Radiological Sciences (J.P.V., N.S., A.B.), and Electrical Engineering (S.M.), University of California, Los Angeles, 300 Stein Plaza, Suite 554E, Los Angeles, CA 90095; and Department of Neurosurgery, University of California, San Francisco, San Francisco, Calif (C. Ames). Received September 6, 2018; revision requested October 15; revision received January 8, 2019; accepted January 14. **Address correspondence to** B.G. (e-mail: *bilwaj@gmail.com*).

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

*Radiology: Artificial Intelligence* 2019; 1(2):e180037 • https://doi.org/10.1148/ryai.2019180037 • Content codes: IN MK

***Purpose:*** To use machine learning tools and leverage big data informatics to statistically model the variation in the area of lumbar neural foramina in a large asymptomatic population.

***Materials and Methods:*** By using an electronic health record and imaging archive, lumbar MRI studies in 645 male (mean age, 50.07 years) and 511 female (mean age, 48.23 years) patients between 20 and 80 years old were identified. Machine learning algorithms were used to delineate lumbar neural foramina autonomously and measure their areas. The relationship between neural foraminal area and patient age, sex, and height was studied by using multivariable linear regression.

***Results:*** Neural foraminal areas correlated directly with patient height and inversely with patient age. The associations involved were statistically significant ($P < .01$).

***Conclusion:*** By using machine learning and big data techniques, a linear model encoding variation in lumbar neural foraminal areas in asymptomatic individuals has been established. This model can be used to make quantitative assessments of neural foraminal areas in patients by comparing them to the age-, sex-, and height-adjusted population averages.



*Supplemental material is available for this article.*

Neural foramina are critical structures in the spine that contain exiting nerve roots. Narrowing of these foramina leads to compression of nerve roots, which results in the clinical syndrome of radiculopathy (1). This syndrome affects more than 200 000 adults in the United States and is a major cause of morbidity and disability (2,3).

Evaluation of radiculopathy is performed by using MRI. In current clinical practice, radiologists interpret these studies and label each foramen as mild, moderately, or severely narrowed or stenosed. These interpretations are critical as they ultimately drive treatment and surgical decisions. However, this practice has been shown to be highly subjective with significant interrater variability (4–7). Moreover, physicians do not currently have a standardized quantitative definition of what a "normal" lumbar neural foraminal area is. Without a proper understanding of what normal is, the variability and subjectivity of abnormal becomes even greater.

To bring objectivity to the diagnosis of neural foraminal narrowing, two key elements are necessary. First, a quantitative approach to delineate neural foramina on MR images and measure their respective areas is required. Ideally, the approach used should be repeatable. That is, when the act of measuring a foramen is repeated by using a particular examination, the exact same value of a neural foraminal area of interest must be obtained. Second, population-based modeling of the variation of neural foraminal areas with demographic factors such as age, sex, and body height must be available.

The purpose of this study was to show that measurement, identification, and delineation of neural foramina by using machine learning–based methods are achievable and useful. Studies on automated segmentation methods based on machine learning have been published in the literature (8–11) and applied to several other regions of the spine (9,10,12–16). In this study, we validated our proposed automated machine learning methods for the delineation of neural foramina on digital MR images. Furthermore, we used our technique to measure foraminal areas in asymptomatic patients and used such measurements to test the hypothesis that these areas change significantly with age, sex, and height. We hope this work will introduce the community to the possibility of a quantitative era of radiologic reporting for spinal stenosis (17).

## Materials and Methods

This study was conducted according to the rules and regulations of the University of California, Los Angeles, as

## Summary

Lumbar neural foramina change in cross-sectional area in the healthy population because of factors such as age, sex, and height, which has important implications for the diagnosis and treatment of spinal stenosis and related degenerative pathologic conditions.

## Key Points

- Machine learning was used to automatically segment neural foramina on 1156 normal spine MR images.
- Foraminal areas were statistically significantly ($P < .01$) correlated with patient age and height.


approved by the institutional review board (IRB #16–000196) with waived consent. This study was Health Insurance Portability and Accountability Act compliant.

### Data Collection

We queried the picture archiving and communication system for individuals who had undergone any form of spine imaging between July 1, 2013, and July 1, 2016, by using the corresponding Current Procedural Terminology (18) codes listed in Table E1 (supplement). This yielded 39 295 unique medical record numbers of those who had undergone either MRI or CT imaging and corresponding accession numbers. We extracted and anonymized the images corresponding to each accession number. Subsequently, we cross-referenced image accession-numbers with anonymized patient records and eliminated studies corresponding to patients whose records were associated with any International Classification of Diseases, Ninth Revision (19), code related to the presence of spinal pathologic conditions or symptoms attributable (eg, sciatica) to the spine. The International Classification of Diseases, Ninth Revision, codes used to filter these studies are listed in Table E2 (supplement). The filtered dataset contained 3837 unique medical record numbers. Although images corresponding to all 3837 medical record numbers were available, we could not use all of these for analyses because of various factors including the unavailability of associated demographic data (670 cases) and failure of the automated analysis to pass quality checks listed in Table 1 (2011 cases). After all filters were applied, we performed our final analysis by using a dataset of 1156 unique patients with an age range of 20–80 years and containing 645 male and 511 female patients. The mean age of men was 50.07 years and that of women was 48.23 years in this study. The average height of men in the study was 69.97 cm and that of women was 64.69 cm.

### MR Sequences

All images used in this study were sagittal T2-weighted MRI acquisitions. The acquisition parameters varied among the patients with mean repetition time of 3756 msec ± 738 (standard deviation) and mean echo time of 107 msec ± 12. The MR acquisitions used for the analysis were three-dimensional acquisitions with a higher in-plane resolution in the sagittal plane. Resolutions in the sagittal plane varied between 0.5 × 0.5 mm per pixel to 2 × 2 mm per pixel. The resolutions were lower in the direction perpendicular to the sagittal plane and ranged from 1 to 5 mm.

**Table 1: Basic Quality Checks**

| Quality Control Questionnaire |
|---|
| Are all five foramina segmented on both sides? |
| Are segmentations > 5 pixels (disks and foramina)? |
| Are disk segmentations and foraminal segmentations adjacent in maximum intensity projection of segmented image? |
| Are there at least five lumbar disks segmented? |
| Does each disk when dilated with a 3 × 3 kernel intersect with at least two neural foramina? Do these foramina intersect? |
| Is any foraminal area as measured < 5 pixels? |

### Training Data Generation for Machine Learning Models

A subset of 100 sagittal MR images were randomly chosen and set aside for algorithmic training purposes. Physicians segmented neural foramina and disks with the help of students. Two independent students (D.V. and C. Ahn) were first trained by physicians (J.B. and L.M.) to find lumbar disks and foramina and delineate them by using ITK-SNAP (20). The students went through each of the 100 three-dimensional sagittal MR images section by section and manually delineated neural foramen and disks. Once the student had delineated the disks and neural foramina, the segmentations were saved as Neuroimaging Informatics Technology Initiative (or NIfTI) files. The imaging data, as well as the segmented NIfTI files, were then reviewed by the physician, who went through all 100 cases on a section-by-section basis and made corrections to the student-generated delineations, when necessary. Thus, we had two sets of neural foraminal (and disk) segmentations which were used strictly for evaluation of interrater variability. We also had a different student (M.E.) segment neural foramina and disks on 100 additional sagittal MR images. However, this student was paired with an attending physician (L.M.), who corrected the resulting segmentations. The dataset segmented by the attending physician was used for training our machine learning algorithms for both disk and neural foraminal segmentations. The number of 100 cases for validation, as well as training, was chosen empirically. As such, we had a limited computational capacity for training deep neural networks and exceeding this number would have made the training process computationally infeasible.

### Training, Validation, and Segmentation with Machine Learning Models

*Neural foraminal segmentation.*—Before segmentation of foramina, sagittal MR images were resampled to 256 × 256 pixels. We used a two-step approach to autonomously segment neural foramina. The first step was the detection step, and the second was the segmentation step. In the detection step, we detected a 25 × 25 pixel window containing the neural foramen. A support vector machine–based object detection system was trained by using histogram of

oriented gradients (21) features and the hard negative mining paradigm to classify if a particular 25 × 25 pixel window contained a foramen. In the segmentation step, an ensemble of regression trees–based (22) shape regression model was used to delineate the foramen. Both the detection step using support vector machines and the delineation step using ensemble of regression trees were implemented by using the Dlib software library.

*Disk segmentation.*—Compared with neural foramina, intervertebral disks are larger anatomic objects. A Deep-U-Net (8) model was trained to segment disks on MR images. The Deep-U-Net we used was implemented by using the Keras application programming interface to the TensorFlow library. The U-Net used here contained four downsampling units, followed by a core unit, followed by four upsampling units, with skip connections between the respective downsampling and upsampling units. Each downsampling unit consisted of a convolutional layer and a max-pooling layer. The convolutional layers had a stride of 1 pixel and a kernel size of 3 × 3. The number of channels went from 32 in the first downsampling layer to 64 in the second one, 128 in the third one, and 256 in the fourth one. The number of inputs went from 256 × 256 in layer 1 to 32 × 32 in layer 4. The core unit was convolutional with 512 channels. The upsampling units consisted of convolutional layers followed by upsampling layers to drive output sizes back from 32 × 32 to 256 × 256, which produced the final output. Rectified linear unit activation was used for convolutional neurons throughout the architecture, except the final output layer, which used sigmoidal activation. The Deep-U-Net model provided a highly accurate detection and segmentation of intervertebral disks.

### Cross-referencing Disk Segmentations with Neural Foraminal Segmentations for Quality Control

Neural foraminal segmentation is substantially more challenging than disk segmentation for both computers and humans. Neural foramina are difficult to segment because they are small and not always bilaterally symmetric. Because of their relatively smaller size, the angle at which they are captured by an MR imager varies due to variations in patient anatomy and position in the imager bore.

Because we know that disk segmentations generated by machine tend to be highly accurate, we could potentially use them as a quality check of the foraminal segmentations. For instance, disks can be labeled from L1 to L5 along the *y*-axis of the image reliably. Depending on the position of patient and image resolution, sometimes a foramen is captured only partially on one side. In such a case, the machine segments on one side and is unable to infer the presence of the foramen on the other side. The disk is conspicuous and continues to be segmented, despite variation in position and anatomy. Because we expect the L1 disk (L1-2 disk) and the two associated foramina to lie adjacent to each other, if two foramina are not found adjacent to the disk, we can conveniently delete this case from the dataset.

This cross-referencing allowed us to eliminate many cases where neural foraminal segmentation was subpar and use the remaining data for exploring the relationships between age, sex, and height and neural foraminal areas. It was perhaps the most important quality control criterion among all the criteria listed in Table 1.

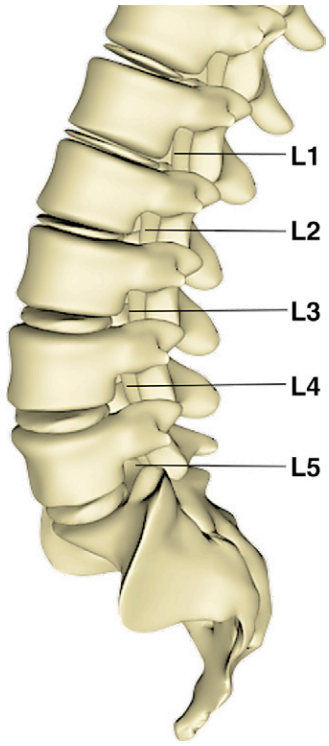


**Figure 1:** Anatomic depiction of lumbar neural foramina.

### Extracted Measurements

Figure 1 depicts lumbar neural foramina that were measured. A particular neural foramen may be captured on more than one sagittal section. If so, we selected the section with the largest cross-sectional area for measurement. We calculated areas separately for the left and right neural foramina at the L1 through L5 levels. The variation of neural foraminal areas as a function of age, sex, and height was subsequently analyzed.

## Results

### Qualitative Results

Figure 2 shows segmentations of neural foramina generated by the machine learning algorithm. Figure 2, *A* and *D,* have left laterality, whereas Figure 2, *B* and *C*, have right laterality. The colors indicate that the computer can distinguish neural foraminal labeling based on the level at which the measurement is made. That is, it can identify L1 as separate or distinct from L2, and so on, and compute neural foraminal areas thereof.

### Quantitative Results

*Validating segmentations.*—Segmentations generated by both the U-Net and the traditional support vector machine–based models were qualitatively and quantitatively assessed. Quantitative assessment was performed by using the Dice coefficient of overlap (23). These results are shown in Table 2.

*Statistical analysis of neural foraminal areas.*—Linear regression analysis (24) was performed by using RStudio (RStudio, Boston, Mass) to analyze the variation of neural foraminal areas, both left and right with respect to age, sex, and height. Sex was coded as 1 for men and 0 for women, and height was measured in centimeters. We explicitly accounted for the interaction between height and sex. The final

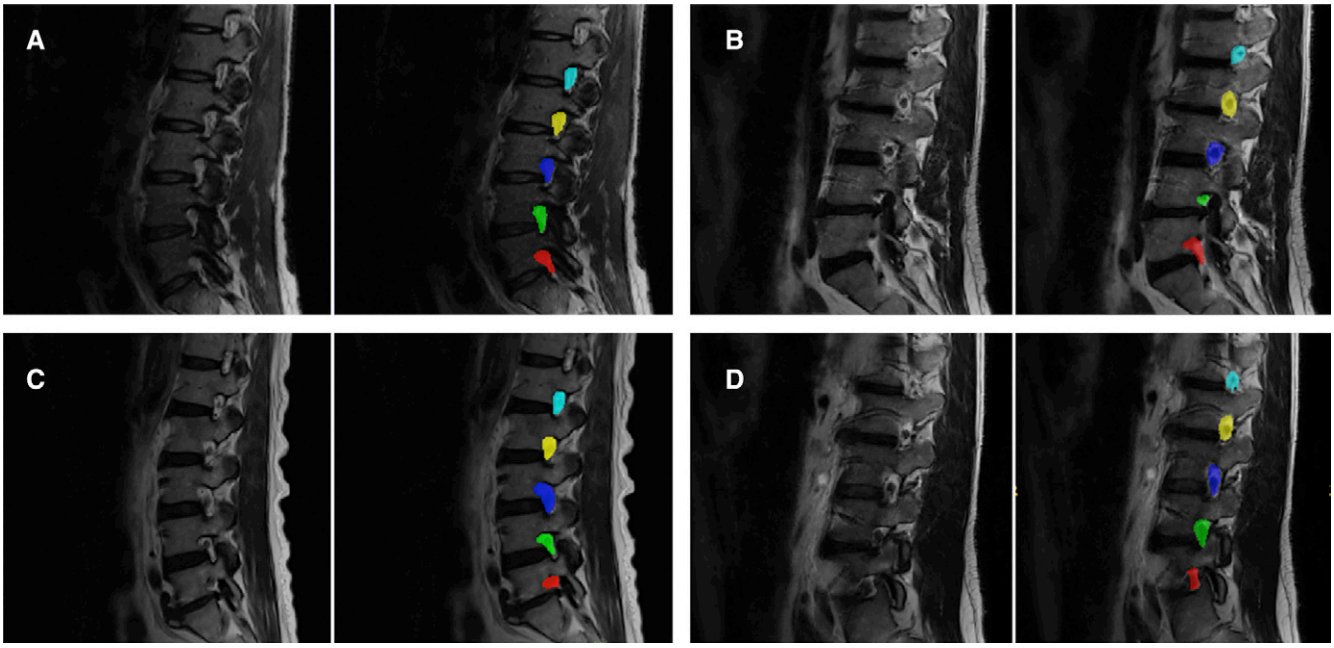


**Figure 2:** *A–D,* A random sample of automated neural foraminal segmentations used for generating measurements. Original MR images (left) and overlaid computer-generated segmentation (right).

multivariable linear regression model was characterized by the equation:

$$\text{Neural foraminal area} = w_1(\text{age}) + w_2(\text{sex}) + w_3(\text{height}) + w_4(\text{height} \cdot \text{sex}) + c,$$

which was fitted to the data. $c$ is the intercept in the regression model. We analyzed whether each coefficient was significantly related to the output variable, namely, the neural foraminal area. The results are tabulated in Table 3. We found that age as well as height were significant predictors of neural foraminal areas. Neural foraminal areas correlated directly with patient height and inversely with patient age. We have assessed significance at $\alpha$ equals .05, and the associated $P$ values and coefficients are shown in Table 3. $P$ values corresponding to a Bonferroni multiple comparisons correction may be obtained from those shown in Table 3 by multiplying by four. This had no effect on the stated results at $\alpha$ equals .05. We also conducted $t$ tests to check if neural foraminal areas measured on the left differed significantly from those measured on the right. The results of these are shown in Table 4.

**Table 2: Quantitative Assessment of Automated Segmentation Algorithms**

| Parameter | Disk | Foramen |
|---|---|---|
| Interrater agreement | 0.91 ± 0.03 | 0.68 ± 0.12 |
| Mean agreement between automated method and humans on test set | 0.84 ± 0.056 | 0.63 ± 0.12 |

Note.—Data are means ± standard deviations.

## Discussion

Machine learning algorithms have induced a paradigm shift in addressing automation and computer-assisted diagnosis (25,26). Given the necessary training datasets, these algorithms can learn to autonomously caption images (27,28), identify objects (29,30), drive cars (31), and, in our case, identify, segment, and measure lumbar neural foramina. In asymptomatic individuals, we found a statistically significant trend of reduced neural foraminal areas with age. We also discovered that taller individuals had larger neural foramina. Furthermore, laterality (left vs right) did not seem to have a statistically significant influence on neural foraminal area.

Algorithms and measurements presented in this article could provide the basis for developing quantitative radiologic reports of the future. Given a model of lumbar neural foraminal areas, as measured by a specific algorithm, abnormalities can be quantified based on the exact degree of deviation from the age-, sex-, and height-specific norms. Thus, in the future, we envision radiologic reports as follows: "Left L3-4 foramen is 2 standard deviations (45%) narrower compared with the population mean corresponding to the patient's age, sex, and height." This approach is diametrically different from the current subjective labeling of mild, moderate, or severe, which is not only variable, but more importantly, does not contain a standardized definition of a normal lumbar neural foramina (without narrowing). We think this computer-aided approach provides a high level of objectivity and reproducibility to radiologic findings, thus representing a major advance compared with the current state of the art. This not only highlights the strength but also the necessity of this radiomics approach to the radiologic investigation of degenerative spine disease.

There were a number of strengths of this study that require mention. First, to develop these machine learning algorithms, a large dataset of annotated MRI data was required. Thus far, such datasets have not been available for spine image analysis. Therefore, the dataset constructed for the purposes of this

study represents an important development and resource in and of itself. Compared with prior studies in the literature regarding image analysis of the spine, the scale of the current study is substantially larger. This approach had two advantages. First, it enabled the use of already collected data, and second, with additional work on diseased cases, it presented the possibility of real-world application in the future.

Although we have established a quantitative paradigm for future radiologic analysis of lumbar spine MR images, this article would be incomplete without acknowledging several limitations and opportunities for future work. The most important limitation was that we have focused exclusively on the analysis of area with respect to lumbar neural foramina. Although this is the most intuitive measure of stenosis or narrowing, we realize that lumbar neural foramina present with substantial shape variation across the population. In fact, this shape variation presented a substantial difficulty in the initial stages of algorithmic development. Perhaps shape quantification along with area will yield additional quantitative measures valuable to both radiologic diagnosis and treatment selection. This remains a topic for future investigation. A second limitation was that some healthy individuals may have an extra lumbar vertebral body (L6), which has not been accounted for in the analysis presented.

Finally, the algorithm presented in this study does not render a diagnosis. The methodology simply identifies, segments, and measures lumbar neural foramina. Through application of this algorithm on a large population of lumbar MR images, we were able to construct a model of what the normal area for a given neural foramina should be in an asymptomatic individual. This is important, especially because age-related spinal degeneration is known to occur in asymptomatic individuals, as well as symptomatic ones (32).

In the future, we will need to perform a similar analysis on patients with clinical symptoms of lumbar radiculopathy to define the abnormal and symptomatic range of areas for lumbar neural foramina and compare these results with radiologist findings contained within the imaging report. Presently, our data only support what the algorithms can identify and measure as neural foraminal area. Nonetheless, challenges remain as the developed methodology would sometimes fail in places where human raters would inevitably succeed.

**Table 3: Effect of Age, Sex, and Height on Foraminal Area as Measured on MR Image by Using a Linear Model**

| Anatomic Region | Age | Sex | Height | Height · Sex |
|---|---|---|---|---|
| Left side | | | | |
| L5 foramen | −0.19 (.00369) | 58.26 (.242) | 2.71 (<.001) | −0.955 (.199) |
| L4 foramen | −0.28 (<.001) | 6.92 (.889) | 2.49 (<.001) | −0.155 (.833) |
| L3 foramen | −0.46 (<.001) | 33.26 (.492) | 3.08 (<.001) | −0.61 (.399) |
| L2 foramen | −0.33 (<.001) | −31.04 (.483) | 2.62 (<.001) | 0.3233 (.625) |
| L1 foramen | −0.37 (<.001) | −9.99 (.820) | 2.49 (<.001) | 0.0023 (.997) |
| Right side | | | | |
| L5 foramen | −0.25 (<.001) | 73.81 (.158) | 2.22 (<.001) | −1.076 (.1696) |
| L4 foramen | −0.33 (<.001) | 4.44 (.929) | 2.00 (<.001) | −0.104 (.888) |
| L3 foramen | −0.41 (<.001) | 80.19 (.0852) | 3.16 (<.001) | −1.31 (.0605) |
| L2 foramen | −0.36 (<.001) | 33.17 (.772) | 3.01 (<.001) | −0.58 (.359) |
| L1 foramen | −0.37 (<.001) | 56.58 (.180) | 2.70 (<.001) | −0.964 (.127) |

Note.—Data are coefficient estimates, with *P* values in parentheses.

**Table 4: *P* Values Comparing Foraminal Areas between Left and Right Sides in Men and Women**

| Anatomic Region | Left versus Right (Men) | Left versus Right (Women) |
|---|---|---|
| L1 foramen | .326 | .27 |
| L2 foramen | .05 | .002 |
| L3 foramen | .008 | .05 |
| L4 foramen | .99 | .47 |
| L5 foramen | .98 | .042 |

In conclusion, we have used artificial intelligence to take the first steps toward establishing a normative model for characterization of areas of lumbar neural foramina on MR images. This model can be used to characterize the variation of neural foraminal areas with respect to age, sex, and height. We believe this ability can provide much needed standardization to the field and facilitate objective evaluations of neural foramina in future radiologic reporting. Hence, the described approach and methodology have the potential to usher in an era of quantitative diagnosis and treatment of degenerative spinal disorders.

**Author contributions:** Guarantors of integrity of entire study, B.G., S.M., M.A., D.B., C. Ames, N.S., L.M.; 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, B.G., J.B., M.E., S.M., D.B., L.M.; clinical studies, B.G., J.B., D.V., C. Ahn, M.E., S.M., D.B., C. Ames, L.M.; statistical analysis, B.G., M.E., D.B.; and manuscript editing, B.G., J.B., M.E., S.M., M.A., D.B., J.P.V., N.S., A.B., L.M.

**Disclosures of Conflicts of Interest:** **B.G.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: employed by UCLA; author's institution (UCLA) receives many grants from various sources, such as NIH, NSF, and other federal sources and foundations. Other relationships: disclosed no relevant relationships. **J.B.** disclosed no relevant relationships. **D.V.** disclosed no relevant relationships. **C. Ahn** disclosed no relevant relationships. **M.E.** disclosed no relevant relationships. **S.M.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the

present article: employed as graduate student researcher by UCLA. Other relationships: disclosed no relevant relationships. **M.A.** disclosed no relevant relationships. **D.B.** disclosed no relevant relationships. **C. Ames** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: consultant for DePuy Synthes, Medtronic, Stryker, Medicrea, K2M, and Biomet Zimmer; employed by UCSF; receives royalties from Stryker, Biomet Zimmer, DePuy Synthes, Nuvasive, Next Orthosurgical, K2M, and Medicrea; provides research for Titan Spine, DePuy Synthes, and ISSG; is on the editorial board of *Operative Neurosurgery*; receives grant funding from SRS; serves on the executive committee of ISSG; and is director of Global Spine Analytics. Other relationships: disclosed no relevant relationships. **J.P.V.** disclosed no relevant relationships. **N.S.** disclosed no relevant relationships. **A.B.** disclosed no relevant relationships. **L.M.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed no relevant relationships. Other relationships: applied for a patient to protect this work through UCLA.

## References


1. Lurie J, Tomkins-Lane C. Management of lumbar spinal stenosis. BMJ 2016;352:h6234.
2. Chad DA. Lumbar spinal stenosis. Neurol Clin 2007;25(2):407–418.
3. Genevay S, Atlas SJ. Lumbar spinal stenosis. Best Pract Res Clin Rheumatol 2010;24(2):253–265.
4. Bankier AA, Levine D, Halpern EF, Kressel HY. Consensus interpretation in imaging research: is there a better way? Radiology 2010;257(1):14–17.
5. Kamper SJ, Apeldoorn AT, Chiarotto A, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and meta-analysis. BMJ 2015;350:h444.
6. Alpert HR, Hillman BJ. Quality and variability in diagnostic radiology. J Am Coll Radiol 2004;1(2):127–132.
7. Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. Ulster Med J 2012;81(1):3–9.
8. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. Medical Image Computing and Computer-assisted Intervention–MICCAI 2015. Cham, Switzerland: Springer, 2015; 234–241.
9. Gaonkar B, Xia Y, Villaroman DS, et al. Multi-parameter ensemble learning for automated vertebral body segmentation in heterogeneously acquired clinical MR images. IEEE J Transl Eng Health Med 2017;5:1800412.
10. Ji X, Zheng G, Liu L, Ni D. Fully automatic localization and segmentation of intervertebral disc from 3D multi-modality MR images by regression forest and CNN. In: Yao J, Vrtovec T, Zheng G, Frangi A, Glocker B, Li S, eds. Computational Methods and Clinical Applications for Spine Imaging. CSI 2016. Lecture Notes in Computer Science, vol 10182. Cham, Switzerland: Springer, 2016; 92–101.
11. Tu Z, Bai X. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. IEEE Trans Pattern Anal Mach Intell 2010;32(10):1744–1757.
12. Peng Z, Zhong J, Wee W, Lee JH. Automated vertebra detection and segmentation from the whole spine MR images. Conf Proc IEEE Eng Med Biol Soc 2005;3:2527–2530.
13. Aslan MS, Ali A, Rara H, et al. A novel 3D segmentation of vertebral bones from volumetric CT images using graph cuts. In: Bebis G, Boyle R, Parvin B, et al, eds. Advances in Visual Computing. ISVC 2009. Lecture Notes in Computer Science, vol 5876. Berlin, Germany: Springer, 2009; 519–528.
14. Zukić D, Vlasák A, Dukatz T, et al. Segmentation of vertebral bodies in MR images. In: Goesele M, Grosch T, Theisel H, Toennies K, Preim B, eds. Vision, Modeling and Visualization. Geneva, Switzerland: The Eurographics Association, 2012; 135–142.
15. Sekuboyina A, Valentinitsch A, Kirschke JS, Menze BH. A localisation-segmentation approach for multi-label annotation of lumbar vertebrae using deep nets. arXiv [preprint] https://arxiv.org/abs/1703.04347. Posted March 13, 2017. Accessed December 2017.
16. Mirzaalian H, Wels M, Heimann T, Kelm BM, Suehling M. Fast and robust 3D vertebra segmentation using statistical shape models. Conf Proc IEEE Eng Med Biol Soc 2013;2013:3379–3382.
17. Hall FM. The radiology report of the future. Radiology 2009;251(2):313–316.
18. Current procedural terminology (CPT). JAMA 1970;212(5):873–874.
19. Centers for Disease Control and Prevention, National Center for Health Statistics. International classification of diseases, ninth revision, clinical modification (ICD-9-CM). Classification of diseases, functioning, and disability. http://www.cdc.gov/nchs/icd/icd9cm.htm. Accessed December 2017.
20. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31(3):1116–1128.
21. Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Piscataway, NJ: IEEE, 2005; 886–893.
22. Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2014; 1867–1874.
23. Gaonkar B, Macyszyn L, Bilello M, et al. Automated tumor volumetry using computer-aided image segmentation. Acad Radiol 2015;22(5):653–661.
24. Rice J. Mathematical statistics and data analysis. Boston: Cengage Learning, 2006.
25. Bishop CM. Pattern recognition and machine learning. New York: Springer. 2006.
26. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–444.
27. Soh M. Learning CNN-LSTM architectures for image caption generation. https://cs224d.stanford.edu/reports/msoh.pdf. Posted 2016. Accessed November 2017.
28. Vinyals O, Toshev A, Bengio S, Erhan D. Show and tell: a neural image caption generator. arXiv [preprint] https://arxiv.org/abs/1411.4555. Posted November 17, 2014. Revised April 20, 2015. Accessed November 2017.
29. Redmon J, Farhadi A. YOLO9000: better, faster, stronger. arXiv [preprint] http://arxiv.org/abs/1612.08242. Posted December 25, 2016. Accessed December 2017.
30. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems 25 (NIPS 2012). https://dl.acm.org/citation.cfm?id=2999257. Published 2012. Accessed December 2017.
31. Chen Z, Huang X. End-to-end learning for lane keeping of self-driving cars. In: 2017 IEEE Intelligent Vehicles Symposium (IV). Piscataway, NJ: IEEE, 2017; 1856–1860.
32. Brinjikji W, Luetmer PH, Comstock B, et al. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol 2015;36(4):811–816.