Title
Vertebral Body Compression Fractures and Bone Density: Automated Detection and Classification on CT Images

Permalink
https://escholarship.org/uc/item/0dt6r9bb

Journal
RADIOLOGY, 284(3)

ISSN
0033-8419

Authors
Burns, JE
Yao, J
Summers, RM

Publication Date
2017-09-01

DOI
10.1148/radiol.2017162100

License
CC BY 4.0

Peer reviewed
Vertebral Body Compression Fractures and Bone Density: Automated Detection and Classification on CT Images

Purpose:
To create and validate a computer system with which to detect, localize, and classify compression fractures and measure bone density of thoracic and lumbar vertebral bodies on computed tomographic (CT) images.

Materials and Methods:
Institutional review board approval was obtained, and informed consent was waived in this HIPAA-compliant retrospective study. A CT study set of 150 patients (mean age, 73 years; age range, 55–96 years; 92 women, 58 men) with (n = 75) and without (n = 75) compression fractures was assembled. All case patients were age and sex matched with control subjects. A total of 210 thoracic and lumbar vertebrae showed compression fractures and were electronically marked and classified by a radiologist. Prototype fully automated spinal segmentation and fracture detection software were then used to analyze the study set. System performance was evaluated with free-response receiver operating characteristic analysis.

Results:
Sensitivity for detection or localization of compression fractures was 95.7% (201 of 210; 95% confidence interval [CI]: 87.0%, 98.9%), with a false-positive rate of 0.29 per patient. Additionally, sensitivity was 98.7% and specificity was 77.3% at case-based receiver operating characteristic curve analysis. Accuracy for classification by Genant type (anterior, middle, or posterior height loss) was 0.95 (107 of 113; 95% CI: 0.89, 0.98), with weighted κ of 0.90 (95% CI: 0.81, 0.99). Accuracy for categorization by Genant height loss grade was 0.68 (77 of 113; 95% CI: 0.59, 0.76), with a weighted κ of 0.59 (95% CI: 0.47, 0.71). The average bone attenuation for T12-L4 vertebrae was 146 HU ± 29 (standard deviation) in case patients and 173 HU ± 42 in control patients; this difference was statistically significant (P < .001).

Conclusion:
An automated machine learning computer system was created to detect, anatomically localize, and categorize vertebral compression fractures at high sensitivity and with a low false-positive rate, as well as to calculate vertebral bone density, on CT images.

Online supplemental material is available for this article.
The societal burden of spine compression fractures is substantial, with an estimated 700,000–750,000 fractures per year in the United States (1–3). Compression fractures are reportedly underdiagnosed due to clinically silent fractures, nonspecific back pain, and unreported fractures (4–7). Early diagnosis is important, as partial compression of one vertebral body portends increased risk for progressive fracturing of that vertebra, compression fracturing of one vertebra positively correlates with subsequent compression fracturing of other vertebrae, and vertebral compression increases the risk profile for hip fractures (3,5,8–12). Mild compression fractures increase patient risk for new fractures, even as detection sensitivity is lower and interobserver variability is higher, lending importance to sensitive and uniform detection (6,13–16). Furthermore, increasing severity of prevalent fractures increases risk for incident fractures; thus, severity grading is important for risk stratification (8,10). Assessment of severity is also important, as compression of more than 50% increases the risk of segmental instability, and compression of more than 80%–90% may obviate vertebroplasty and kyphoplasty (2,17–19). Incident fracture risk is additionally linked to compression morphology, lending importance to determination of shape type (11). Finally, increased risk for vertebral fractures is independently linked to osteoporosis, with the combined risk factors of osteoporosis and prevalent fractures acting as risk multipliers (20–23). Early diagnosis enables one to initiate administration of bisphosphonates and selective estrogen receptor modulators to decrease the risk of subsequent fractures, versus management with an intervention such as kyphoplasty (7,17–19). A computer system that could be used to sensitively detect compression, classify the degree of height loss and pattern of compression, and assess bone mineral density would be useful in clinical detection and classification. This system could be integrated into an opportunistic detection framework to assess CT images obtained for indications other than vertebral fracture. It could also be used in research studies on risk prediction in clinical decision making.

The widely used semiquantitative Genant classification system was developed in the 1990s to assess vertebral compression deformities on two-dimensional radiographic images (3). We have created a system that updates the schema by using the full power of three-dimensional (3D) computed tomographic (CT) images with lateralization of the compression fracture patterns. When coupled with metrics such as bone density, this system has potential for use in prospective clinical trials, in the generation of risk factors for progressive vertebral body compression fracturing, and in the generation of statistical profiles for patients who may benefit from intervention rather than conservative care.

The purpose of this study was to create and validate a computer system with which to detect, localize, and classify compression fractures and to calculate bone density of thoracic and lumbar vertebral bodies on CT images.

### Materials and Methods

#### Study Subjects

Our study was compliant with the Health Insurance Portability and Accountability Act and was conducted with institutional review board approval. Informed consent was waived for this retrospective study in which we analyzed previously obtained images.

#### Advances in Knowledge

- A fully automated machine learning software system with which to detect, localize, and classify compression fractures and determine the bone density of thoracic and lumbar vertebral bodies on CT images was developed and validated.
- The computer system has a sensitivity of 95.7% in the detection of compression fractures and in the localization of these fractures to the correct vertebrae, with a false-positive rate of 0.29 per patient.
- The accuracy of this computer system in fracture classification by Genant type was 95% (weighted \( \kappa = 0.90 \)).

#### Implications for Patient Care

- This proof-of-function computer system depicts compression fractures of the thoracic and lumbar vertebrae with a high degree of sensitivity, devoid of interobserver variability, potentially enabling early detection and treatment to mitigate the risk of fracture progression or additional fractures.
- This system enables one to discern the severity and pattern of compression of the fractured vertebra and bone density, and it has potential for use in risk stratification for further fracturing to help guide patient care.
- Potential benefits of this system include automated classification of fractures with an extended three-dimensional Genant classification schema and standardized reporting of compression fractures.
Reports from CT examinations performed with the spine protocol were drawn from the imaging database of the University of California-Irvine medical center by using the radiology information system application Radiology Report Search (RadNet; Cerner Millennium, North Kansas City, Mo). The date of service for examinations selected ranged from 2012 to 2015. Figure E1 (online) shows a Standards for Reporting Diagnostic Accuracy Studies chart for the report search method.

Reports dictated as positive for thoracic or lumbar compression fractures within the search region from the T1 to L5 vertebrae were flagged, and associated CT images were reviewed with an Impax picture archiving and communications system, (AGFA, Mortsel, Belgium). During this review, exclusion criteria (Fig E1 [online]) for case selection were applied. A case set of 75 consecutive patients with one or more compression fractures was assembled. Additionally, images in 75 patients without fractures who were matched by age and sex to patients in the case set were assembled to constitute the control set. The mean age of patients was 73 years ± 11 (standard deviation), with an age range of 55–96 years. There were 46 women and 29 men in both the patient group and the group of age- and sex-matched control subjects. In the female cohort, the mean age was 76 years ± 10 (range, 56–92 years). In the male cohort, the mean age was 69 years ± 11 (range, 55–96 years). The age difference between the male and female cohorts was significant ($P < .001$).

Compression Fracture Identification and Description

CT images from the picture archiving and communication system were saved with soft-tissue reconstruction kernel in uncompressed Digital Imaging and Communications in Medicine format. Section thickness of the images ranged from 1.0 to 2.0 mm, and in-plane resolution ranged from 0.31 to 0.45 mm. In the 75-patient CT case set, the approximate centroid of each of the 210 thoracic and lumbar vertebrae that showed vertebral body compression fracture deformities was electronically marked and recorded (J.E.B., a board-certified radiologist with 10 years of experience).

Additionally, an enhanced Genant classification of vertebral body compression fractures by type (wedge, concave, or crush) and grade of height loss (grade 1, <25%; grade 2, 26%–40%; grade 3, >40%) was performed for 113 vertebrae in the case set. Classification was restricted to studies with no more than two contiguous fracture levels to mitigate the percentage height loss scaling inaccuracies in the assessment of patients with multiple contiguous compressed levels due to the lack of a standardized uniform anatomically proximate comparator. In creating the reference standard data set for percentage height loss, the reference point used was the nearest normal-appearing vertebra, which was chosen with qualitative visual inspection.

The axial profile of each vertebral body was divided into five zones (Fig 1). This zone division was designed to conform to the original Genant classification schema, with allowance for lateralization and development of an enhanced Genant classification schema. The two lateral zones created by the sagittal plane partitions are defined as $R$ (right) and $L$ (left). Three zones were created by partitioning the central sagittal region into three central coronal zones $A$ (anterior), $C$ (central), and $P$ (posterior).

Figure 1: Axial CT image of a vertebra shows zone division of the vertebral body. The axial section of the vertebral bodies was divided into zones for determination of subsection of height loss. This zone division was designed to conform to the original Genant classification schema, with allowance for lateralization and development of an enhanced Genant classification schema. The two lateral zones created by the sagittal plane partitions are defined as $R$ (right) and $L$ (left). Three zones were created by partitioning the central sagittal region into three central coronal zones $A$ (anterior), $C$ (central), and $P$ (posterior).
The characteristics of the fractures are then graded by using a support vector regression technique, which is a supervised machine learning technique that associates vertebral height distribution with the fracture grades and types.

Fully automated bone densitometry computation was then performed for T12-L4 vertebrae. Bone density algorithms had an origin in previous software developed to determine bone mineral density in the spine at CT (26). Vertebral marked positive with the fracture-detection algorithms were excluded from densitometry calculations because of alterations in native bone density inherent in the compressive process. With previously computed vertebral segmentation and partitioning, the cortex was removed from the underlying trabecular bone via a process of adaptive erosion by up to 5 mm of the vertebral body periphery. Bone density was then estimated by averaging the attenuation (in Hounsfield units) within the remaining 3D medullary space (26, 27).

Analysis of each case lasted approximately 5 minutes and was performed with a high-performance desktop computer with a dual 2.30-GHz central processing unit, 16.0-GB memory, and a 64-bit Windows 7 operating system (Dell Precision T7600; Dell, Round Rock, Tex).

The false-positive and false-negative compression fracture detections were reviewed and were decided in consensus by two board-certified fellowship-trained

The characteristics of the fractures are then graded by using a support vector regression technique, which is a supervised machine learning technique that associates vertebral height distribution with the fracture grades and types.

Fully automated bone densitometry computation was then performed for T12-L4 vertebrae. Bone density algorithms had an origin in previous software developed to determine bone mineral density in the spine at CT (26). Vertebral marked positive with the fracture-detection algorithms were excluded from densitometry calculations because of alterations in native bone density inherent in the compressive process. With previously computed vertebral segmentation and partitioning, the cortex was removed from the underlying trabecular bone via a process of adaptive erosion by up to 5 mm of the vertebral body periphery. Bone density was then estimated by averaging the attenuation (in Hounsfield units) within the remaining 3D medullary space (26, 27).

Analysis of each case lasted approximately 5 minutes and was performed with a high-performance desktop computer with a dual 2.30-GHz central processing unit, 16.0-GB memory, and a 64-bit Windows 7 operating system (Dell Precision T7600; Dell, Round Rock, Tex).

The false-positive and false-negative compression fracture detections were reviewed and were decided in consensus by two board-certified fellowship-trained

The characteristics of the fractures are then graded by using a support vector regression technique, which is a supervised machine learning technique that associates vertebral height distribution with the fracture grades and types.

Fully automated bone densitometry computation was then performed for T12-L4 vertebrae. Bone density algorithms had an origin in previous software developed to determine bone mineral density in the spine at CT (26). Vertebral marked positive with the fracture-detection algorithms were excluded from densitometry calculations because of alterations in native bone density inherent in the compressive process. With previously computed vertebral segmentation and partitioning, the cortex was removed from the underlying trabecular bone via a process of adaptive erosion by up to 5 mm of the vertebral body periphery. Bone density was then estimated by averaging the attenuation (in Hounsfield units) within the remaining 3D medullary space (26, 27).

Analysis of each case lasted approximately 5 minutes and was performed with a high-performance desktop computer with a dual 2.30-GHz central processing unit, 16.0-GB memory, and a 64-bit Windows 7 operating system (Dell Precision T7600; Dell, Round Rock, Tex).

The false-positive and false-negative compression fracture detections were reviewed and were decided in consensus by two board-certified fellowship-trained

The characteristics of the fractures are then graded by using a support vector regression technique, which is a supervised machine learning technique that associates vertebral height distribution with the fracture grades and types.

Fully automated bone densitometry computation was then performed for T12-L4 vertebrae. Bone density algorithms had an origin in previous software developed to determine bone mineral density in the spine at CT (26). Vertebral marked positive with the fracture-detection algorithms were excluded from densitometry calculations because of alterations in native bone density inherent in the compressive process. With previously computed vertebral segmentation and partitioning, the cortex was removed from the underlying trabecular bone via a process of adaptive erosion by up to 5 mm of the vertebral body periphery. Bone density was then estimated by averaging the attenuation (in Hounsfield units) within the remaining 3D medullary space (26, 27).

Analysis of each case lasted approximately 5 minutes and was performed with a high-performance desktop computer with a dual 2.30-GHz central processing unit, 16.0-GB memory, and a 64-bit Windows 7 operating system (Dell Precision T7600; Dell, Round Rock, Tex).

The false-positive and false-negative compression fracture detections were reviewed and were decided in consensus by two board-certified fellowship-trained
radiologists (J.E.B., R.M.S.; the latter had 19 years of experience).

Statistical Analyses
System performance for the CT data set was gauged by using free-response receiver operating characteristic (FROC) curve analysis with 10-fold cross validation. When we varied the support vector machine probability threshold in the determination of compression fracture versus a nonfractured vertebra, it resulted in the generation of the FROC curve. Confidence intervals (CIs) were obtained by resampling the test set 1000 times (bootstrapping) and averaging the performance.

The target computer-aided detection sensitivity in compression fracture detection was set at 90% or greater, which was in balance with a clinically reasonable false-positive rate, when we chose the FROC analysis operating point.

Results
Of the 75 patients in the case set, 25 had a fracture in one vertebral body, 19 had a fracture in two vertebral bodies, 13 had fractures in three vertebral bodies, four had fractures in four vertebral bodies, and 14 had fractures in five or more vertebral bodies for a total of 210 fractured thoracic and lumbar vertebral bodies with more than 10% height loss in the case set. An average of 2.8 vertebral body compression fracture deformities per patient and a range of one to nine compressed vertebrae per patient were identified in the case set. A total of 1275 vertebrae were evaluated with the computer system in the combined set of 150 CT examinations (case group, n = 75; control group, n = 75).

Sensitivity in the detection of vertebrae with compression fracture deformities was 95.7% relative to the reference standard (201 of 210 vertebrae, 95% CI: 87.0%, 98.9%), with a false-positive rate of 0.29 per patient (Fig 3). Additionally, at case-based receiver operating characteristic curve analysis, sensitivity was 98.7% and specificity was 77.3%.

Figure 3: FROC curve of support vector machine performance in the detection of vertebral bodies with compression fractures shows 95.7% sensitivity (95% CI: 87%, 98.9%) with a false-positive rate of 0.29 per patient. Error bars represent 95% CIs.

Figure 4: Sagittal CT section shows false-negative findings in compression fracture detection with the computer system in a 77-year-old male patient. Missed concave fractures (arrows) are visible at T8 and T11.
due to osteopenia, and one (2%) was due to Scheuerman disease. Seven (16%) of the initially categorized false-positive findings were found to be mild compression fractures at secondary review; thus, these were true-positive detections omitted from the reference standard data set.

Fracture detection in the case set showed a bimodal distribution, with local peaks in the number of fractures at the L1 and T7 levels. The largest distribution of fractures was about the thoracolumbar junction (Fig 6).

The most common morphologic compression type found in the case set was the wedge-type anterior compression fracture. There were 80 wedge-type fractures, three crush fractures, and 30 biconcave fractures. The most common grade of fracture was grade 3 (>40% height loss). There were 35 grade 1, 31 grade 2, and 47 grade 3 fractures detected with the system.

System performance was determined relative to the manual annotation data set for the morphology of compression deformities. There was 95% overall agreement (107 of 113 vertebrae, 95% CI: 0.89, 0.98) for categorization of compression morphology (wedge, concave, or crush fracture) with the system relative to radiologist classification (Table 2).

The calculated weighted κ coefficient was 0.90 (95% CI: 0.81, 0.99), which was consistent with the highest level agreement category (ie, very good agreement). FROC curves for compression morphology are shown in Figure E4 (online); crush-type morphology was not included in the curves because of small statistics.

System performance was also determined relative to the manual annotation data set for the severity grade of compression deformities. There was 68% overall agreement (77 of 113 vertebrae, 95% CI: 0.59, 0.76) for categorization of compression severity (grade 1, 2, or 3)
Table 2

| Confusion Matrix of Genant Fracture Morphology: Radiologist versus Computer Assessment of Fracture Type | Computer Classification |
|---|---|---|
| Radiologist Classification | Wedge Fracture | Concave Fracture | Crush Fracture |
| Wedge fracture | 80 | 0 | 0 |
| Concave fracture | 5 | 25 | 0 |
| Crush fracture | 0 | 0 | 3 |

Note.—Data are number of cases with concordance (diagonal values) versus discrepancy (off-diagonal values) in the classification of Genant fracture morphology by radiologist and by computer system. Of the 210 fractured vertebrae, 113 were assessed for type and grade of height loss and are listed in this table.

Table 3

| Confusion Matrix of Genant Severity Grade: Radiologist versus Computer Assessment |
|---|---|---|
| Radiologist Grade | Computer Grade 1 | Grade 2 | Grade 3 |
| Grade 1 | 23 | 10 | 2 |
| Grade 2 | 3 | 26 | 2 |
| Grade 3 | 2 | 17 | 28 |

Note.—Data are number of cases with concordance (diagonal values) versus discrepancy (off-diagonal values) in the categorization of Genant severity grade by radiologist and by computer system. Genant grades of height loss are as follows: grade 0, less than 15% height loss; grade 1, 15%–25% height loss; grade 2, 26%–40% height loss; and grade 3 more than 40% height loss. Of the 210 fractured vertebrae, 113 were assessed for type and grade of height loss and are listed in this table.

with the system relative to radiologist classification (Table 3). The calculated weighted $\kappa$ coefficient was 0.59 (95% CI: 0.47, 0.71), which was indicative of moderate or good agreement. FROC curves of system performance for compression severity are shown in Figure 7. Additionally, system performance for severity grade of compression was also determined relative to manual annotation for each of the Genant-enhanced classification zones (A, C, P, R, or L). There was 59% overall agreement (66 of 113, 95% CI: 0.49, 0.67) for categorization of compression severity for zone A, 65% overall agreement (65 of 113, 95% CI: 0.49, 0.67) for categorization of compression severity for zone C, and 86% overall agreement (97 of 113, 95% CI: 0.78, 0.91) for categorization of compression severity for zone P. There was 69% overall agreement (78 of 113, 95% CI: 0.60, 0.77) for categorization of compression severity for zone B and 69% overall agreement (78 of 113, 95% CI: 0.60, 0.77) for categorization of compression severity for zone L. Figure 5 shows examples of compression fracture grading by a radiologist and with a computer. Additional examples of compression fracture grading can be found in Figure E5 (online).

Bone densitometry results showed average bone density in patients with fractures was less than that in those without fractures, as we intuitively expected (Table 4). The average bone density for the T12-L4 interval was 146 HU ± 29 in the case set and 173 HU ± 42 in the control set. The difference in bone density was significant between case and control sets, both at individual vertebral levels from T12 to L4 as well as in summation ($P < .001$, paired $t$ test). On an individual basis, bone density of each age- and sex-matched control patient exceeded that of the corresponding case patient 67%–73% of the time, depending on the level from T12 to L4.

Discussion

We designed and validated an automated machine learning system to detect and enumerate the level of vertebral compression fractures and to classify fracture morphology and severity according to Genant standards. We have updated the Genant schema to use the full power of 3D CT images. Additionally, this system was designed to delineate bone density, with the potential to provide combined metrics for risk stratification of progressive compression fracturing of prevalent vertebral and incident vertebral fracturing based on both shape and compositional feature metrics.

The system holds potential to decrease interobserver variability of Genant classification via systematic and uniform quantitative compression assessment. The vertebral body is a 3D structure, and the system can be used to evaluate asymmetric left- and right-sided compression, extending the two-dimensional radiographic Genant system into three dimensions. As such, the system also enables assessment of postcompression scoliosis and kyphosis. This system holds potential in the determination of developmentally related scoliosis, and Scheuermann disease.

This system performed with 95.7% sensitivity in fracture detection and localization to the correct vertebral level, with a low false-positive rate. There was a high level of overall agreement (95%) for compression morphology and 68% overall agreement for severity categorization relative to radiologist classification.

Quantitative data generated by our system have the potential to decrease interobserver variability in Genant classification via systematic and uniform quantitative compression assessment. Detailed information regarding morphology and severity, combined with the per-patient speed of analysis (approximately 5 minutes), can be used to predict future ability to provide repeatable, timely, and detailed evaluation of compression patterns within the Genant schema and to allow extension to new 3D classification schemas. Fracture risk stratification is of current interest in patients with osteopenia or osteoporosis (28). The stacked height compass model forms a global height loss view of the spinal column and may be applicable to the semiquantitative spinal deformity index risk matrix (29). Quantitative morphologic features of vertebrae, both
Vertebral body compression fractures. Normal architecture of the adjacent vertebrae can be variable due to degenerative change, variant anatomy, and compression fracture, limiting its use as a reference point. A normal increase in vertebral body size proceeding from cranial to caudal incorporates systematic error into the measure of relative height loss. The computer determined percentage height loss by uniform and systematic averaging of multiple adjacent vertebrae; this was difficult to replicate individually and statistically as a group, in combination with bone density generated with the system, may further aid comparative effectiveness research, thereby generating predictive features and risk profiles for further fracturing and helping to guide the development of new treatment paradigms based on evidence-based medicine patient outcome studies.

The bimodal distribution of fractures mirrored that which was expected and reported in other works (15,21). Fracture severity was somewhat unusual. Grade 3 deformities were the most common finding, whereas we had expected grade 1 deformities to be the most common finding on the basis of prior studies (30). Possible reasons for this discrepancy include small sample size and sampling bias.

As intuitively expected, the system calculated a significantly lower bone density in the case population than in the age- and sex-matched control subjects, on average. On an individual case-matched basis, bone density of patients in the case group exceeded that of patients in the control group 27%–33% of the time. This was thought to be most likely related to selection bias and variation in the individual medical histories of patients in this small sample.

Prior work to assess vertebral compression fractures on lateral radiographs has reached the stage of clinical application (31,32). We have designed a system to assess height loss on CT images and to incorporate the 3D characteristics of the vertebral body. When compared with the method of Hsieh et al (33), vertebral body height loss is measured directly with 3D segmentation and partitioning of the vertebrae rather than indirectly by measuring spinal curvature. We have incorporated the left and right aspects for 3D lateralization, extending beyond anterior, middle, and posterior sagittal midline vertebral body height loss as in Ghosh et al (34) and Baum et al (25), and compared with Baum et al (25), we include the entire thoracic spine, not just the T5-T12 vertebrae. Additionally, we measured bone density in this study. Our group performed preliminary work in direct determination of height loss in the height compass model, without creation of a Genant model and without Genant categorization or assessment of bone density (35).

The main limitation of this system is thought to be the lack of concrete comparative manual reference standards with which to determine the percentage height loss. Posterior body wall compression can limit use as a reference standard for anterior compression. Normal architecture of the adjacent vertebrae can be variable due to degenerative change, variant anatomy, and compression fracture, limiting its use as a reference point. A normal increase in vertebral body size proceeding from cranial to caudal incorporates systematic error into the measure of relative height loss. The computer determined percentage height loss by uniform and systematic averaging of multiple adjacent vertebrae; this was difficult to replicate individually and statistically as a group, in combination with bone density generated with the system, may further aid comparative effectiveness research, thereby generating predictive features and risk profiles for further fracturing and helping to guide the development of new treatment paradigms based on evidence-based medicine patient outcome studies.

As intuitively expected, the system calculated a significantly lower bone density in the case population than in the age- and sex-matched control subjects, on average. On an individual case-matched basis, bone density of patients in the case group exceeded that of patients in the control group 27%–33% of the time. This was thought to be most likely related to selection bias and variation in the individual medical histories of patients in this small sample.

Prior work to assess vertebral compression fractures on lateral radiographs has reached the stage of clinical application (31,32). We have designed a system to assess height loss on CT images and to incorporate the 3D characteristics of the vertebral body. When compared with the method of Hsieh et al (33), vertebral body height loss is measured directly with 3D segmentation and partitioning of the vertebrae rather than indirectly by measuring spinal curvature. We have incorporated the left and right aspects for 3D lateralization, extending beyond anterior, middle, and posterior sagittal midline vertebral body height loss as in Ghosh et al (34) and Baum et al (25), and compared with Baum et al (25), we include the entire thoracic spine, not just the T5-T12 vertebrae. Additionally, we measured bone density in this study. Our group performed preliminary work in direct determination of height loss in the height compass model, without creation of a Genant model and without Genant categorization or assessment of bone density (35).

The main limitation of this system is thought to be the lack of concrete comparative manual reference standards with which to determine the percentage height loss. Posterior body wall compression can limit use as a reference standard for anterior compression. Normal architecture of the adjacent vertebrae can be variable due to degenerative change, variant anatomy, and compression fracture, limiting its use as a reference point. A normal increase in vertebral body size proceeding from cranial to caudal incorporates systematic error into the measure of relative height loss. The computer determined percentage height loss by uniform and systematic averaging of multiple adjacent vertebrae; this was difficult to replicate individually and statistically as a group, in combination with bone density generated with the system, may further aid comparative effectiveness research, thereby generating predictive features and risk profiles for further fracturing and helping to guide the development of new treatment paradigms based on evidence-based medicine patient outcome studies.

### Table 4

<table>
<thead>
<tr>
<th>Vertebral Level</th>
<th>Case Attenuation (HU)</th>
<th>Control Attenuation (HU)</th>
<th>Control Attenuation Greater than Case Attenuation (%)</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T12</td>
<td>146 ± 29</td>
<td>169 ± 37</td>
<td>67</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>L1</td>
<td>141 ± 26</td>
<td>173 ± 43</td>
<td>73</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>L2</td>
<td>147 ± 28</td>
<td>171 ± 40</td>
<td>68</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>L3</td>
<td>146 ± 31</td>
<td>172 ± 44</td>
<td>67</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>L4</td>
<td>146 ± 29</td>
<td>179 ± 45</td>
<td>73</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Average</td>
<td>146 ± 29</td>
<td>173 ± 42</td>
<td>71</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Note.—Bone attenuation is given as mean ± standard deviation.
manually and was likely more accurate than manual assessment. A second limitation was related to the finite size of each axial zone with (generally) continuous variation of vertebral body height within each zone. The reference standard data were created by visual inspection of images for the approximate midpoint, allowing for skewed random variation in the choice of reference point. A third limitation was that our estimation of sensitivity was based on use of a cutoff derived from the data and the resulting FROC curve; thus, the calculated 95.7% sensitivity may be an overestimate. Finally, the bone density determined was not normalized into bone mineral density units, as an external reference was not available for all patients in this retrospective study. However, Hounsfield units from clinical CT examinations have been previously shown to correlate with dual energy x-ray absorptiometry values (27).

In conclusion, we have designed and validated a fully automated quantitative system with which to detect and enumerate levels of thoracic and lumbar vertebral body compression fractures on CT images, determine Genant classification of compression severity and morphology, generate extended Genant scores for fracture lateralization, and generate bone densities. Our work is proposed as a proof-of-concept system for fracture risk stratification and clinical decision support.

Disclosures of Conflicts of Interest: J.E.B. disclosed no relevant relationships. J.Y. Activities related to the present article: disclosed no relevant relationships. R.M.S. Activities related to the present article: disclosed no relevant relationships. J.Y. Activities related to the present article: disclosed no relevant relationships. R.M.S. Activities related to the present article: disclosed no relevant relationships. J.E.B. Activities related to the present article: disclosed no relevant relationships. R.M.S. Activities related to the present article: disclosed no relevant relationships.

References

26. Summers RM, Baecher N, Yao J, et al. Feasibility of simultaneous computed tomographic colonography and fully automated bone min-


