

1 Definition of normal vertebral morphology  
2 using NHANES-II radiographs  
3

4 John A. Hipp, PhD<sup>1\*</sup>, Trevor F. Grieco<sup>1</sup>, PhD, Patrick Newman, MS<sup>1</sup>, Charles A. Reitman,  
5 MD<sup>2</sup>

6 <sup>1</sup> Medical Metrics, Imaging Core Laboratory, Houston, TX 77056

7 <sup>2</sup> Orthopaedics and Physical Medicine, Medical University of South Carolina, Charleston, SC 29425  
8  
9

10 \* Corresponding Author: John A. Hipp, PhD

11 Email address: [jhipp@medicalmetrics.com](mailto:jhipp@medicalmetrics.com)

12 Mailing address: 1632 Summit Lake Shore Rd, NW, Olympia, WA 98502

13 Phone: (360) 402-0022  
14  
15

## 1 Abstract

2 **Background:** A robust definition of normal is required to confidently identify vertebral abnormalities  
3 such as fractures. Between 1976 and 1980, the 2nd National Health and Nutrition Examination Survey  
4 (NHANES-II) was conducted. Justified by the prevalence of neck and back pain, approximately 10,000  
5 lateral cervical spine and 7,000 lateral lumbar spine X-rays were collected. Demographic,  
6 anthropometric, health, and medical history data were also collected. This resource can be used for  
7 establishing normative reference data that can subsequently be used to diagnose abnormal vertebral  
8 morphology.

9 **Purpose:** 1) Develop normative reference data for vertebral morphology using the lateral spine  
10 radiographs from NHANES-II. 2) Document sources of variability.

11 **Subject Sample:** Nationwide probability sample to document health status of the United States.

12 **Methods:** The coordinates of the four vertebral body corners were obtained using previously  
13 validated, automated technology consisting of a proprietary pipeline of neural networks and coded  
14 logic. These landmarks were used to calculate six vertebral body morphology metrics: 1)  
15 anterior/posterior vertebral body height ratio (VBHR); 2) superior/inferior endplate width ratio (EPWR);  
16 3) forward/backward diagonal ratio (FBDR); 4) height/width ratio (HWR); 5: angle between endplates  
17 (EPA); 6) Angle between posterior wall and superior endplate (PSA). Descriptive statistics were  
18 generated and used to identify and trim outliers from the data and obtain a gaussian distribution for  
19 each metric. Descriptive statistics were tabulated using the trimmed data for use in quantifying  
20 deviation from average for each metric. The dependency of these metrics on sex, age, race, nation of  
21 origin, height, weight, and BMI was also assessed.

1 **Results:** Computer generated lumbar landmarks were obtained for 42,980 vertebrae from lumbar  
2 radiographs and 54,093 vertebrae from cervical radiographs for subjects 25 to 74 years old. After  
3 removing outliers, means and standard deviations for the remaining 35,275 lumbar and 44,938 cervical  
4 vertebrae changed only slightly, suggesting that normal morphology and intervertebral alignment is  
5 dominant in the data. There was low variation in vertebral morphology after accounting for vertebra (L1,  
6 L2, etc.), and the  $R^2$  was high for analyses of variance. The EPWR, FBDR and PSA generally had the lowest  
7 coefficients of variation. Excluding outliers, Age, sex, race, nation of origin, height, weight, and BMI were  
8 statistically significant for most of the variables, though the F-statistic was very small compared to that  
9 for vertebral level. Excluding all variables except vertebra changed the  $R^2$  very little (e.g. for the lumbar  
10 data, VBHR  $R^2$  went from 0.804 to 0.795 and FBDR  $R^2$  went from 0.9005 to 0.9000). Reference data were  
11 generated that can be used to produce standardized metrics in units of standard deviation from  
12 average. This allows for easy identification of abnormalities resulting from vertebral fractures, atypical  
13 vertebral body morphologies, and other congenital or degenerative conditions. Standardized metrics  
14 also remove the effect of vertebra thereby enabling data for all vertebrae to be pooled in research  
15 studies.

16 **Conclusions:** The NHANES-II collection of spine radiographs and associated data may prove to be a  
17 valuable resource that can facilitate standardized spine metrics useful for objectively identifying  
18 abnormalities. The data may be particularly valuable for identification of vertebral fractures, although X-  
19 rays taken early in life would be needed in some cases to differentiate between normal anatomic  
20 variants, fractures, and vertebral shape remodeling.

21

## 1 Introduction

2           Although vertebral fractures can be asymptomatic, these fractures can be a source of  
3 symptoms. One of the most important predictors of subsequent fractures is the number of prior  
4 fractures.<sup>(1-3)</sup> Vertebral fractures are therefore diagnostically important. Fractures can be missed in  
5 clinical practice<sup>(4,5)</sup>. Ferrar et al. succinctly identified the challenge of fracture detection: “The  
6 identification of vertebral fractures is problematic because (1) “normal” radiological appearances in the  
7 spine vary greatly both among and within individuals; (2) “normal” vertebrae may exhibit misleading  
8 radiological appearances due to radiographic projection error; and (3) “abnormal” appearances due to  
9 non-fracture deformities and normal variants are common, but can be difficult to differentiate from true  
10 vertebral fracture.”<sup>(6)</sup>

11           A typical definition of a fracture is “20% or greater loss in expected vertebral body height.” This  
12 definition requires definition of “expected” vertebral body height and that requires accurate  
13 measurements of vertebral body shape in a large and representative population. Validation of the  
14 clinically meaningful threshold level for change in vertebral dimensions also requires a reliable reference  
15 standard. Inaccuracies in measuring shape and the challenge of finding a suitable reference standard are  
16 well documented.<sup>(6-43)</sup> One option is to use the patient’s own adjacent vertebral bodies as a reference,  
17 but there is no definitive method to assure that the adjacent levels are normal, or reference data to  
18 know how much variability between levels is normal. Normative vertebral body shape metrics that  
19 represent a large proportion of the population and account for age, sex, and other confounding factors  
20 have not been previously published. Robust reference data and an understanding of the important  
21 confounding factors would facilitate more accurate and reproducible diagnosis of the presence and  
22 severity of vertebral body fractures.

1           Between 1976 and 1980, the 2nd National Health and Nutrition Examination Survey (NHANES-II)  
2 was conducted<sup>1</sup>. This was a nationwide probability sample to document health status of the United  
3 States. Justified by the prevalence and societal impact of neck and back pain, approximately 10,000  
4 lateral cervical spine and 7,000 lateral lumbar spine X-rays were collected as part of this survey.  
5 Demographic, anthropometric, health, and medical history data were also collected. This resource can  
6 be used for establishing normative reference data that can subsequently be used to diagnose  
7 abnormalities such as fractures, vertebral deformities, and congenital abnormalities. These data can also  
8 be used to determine if the sex, age, race, height, weight and/or body mass index (BMI) of an individual  
9 are needed to predict normal vertebral morphology.

## 10   Methods

11           The NHANES-II images and data were obtained through public access<sup>2</sup>. A previously validated  
12 series of neural networks and coded logic (Spine Camp™, Medical Metrics, Inc., Houston, TX) were used  
13 to obtain four landmarks for each vertebral body from L1 to S1 (Figure 1). Although the NHANES-II  
14 lumbar X-rays almost always included up to T10 in the field-of-view, the neural networks were only  
15 trained to recognize L1 to S1. The neural networks include quality control networks to assure the X-ray  
16 was a lateral lumbar view, manipulate the image if needed so upper vertebrae are toward the top of the  
17 X-ray, with spinous processes toward the left side, and with bone being whiter than air. Neural networks  
18 and coded logic were then used to segment the bone, find individual vertebrae, label the vertebrae, find  
19 the four corners, and finally refrain from reporting landmark coordinates when the confidence of the  
20 networks was too low. All the neural networks were trained with over 50 thousand lateral X-rays where  
21 analysts had previously digitized standardized landmarks. The NHANES images were not used in training

---

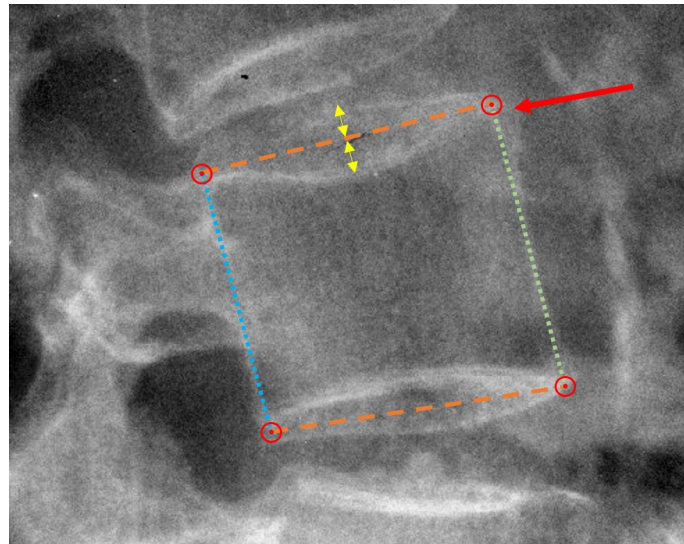
<sup>1</sup> <https://www.cdc.gov/nchs/nhanes/nhanes2>

<sup>2</sup> <https://www.cdc.gov/nchs/nhanes/index.htm>

1 the neural networks. The analysts used Quantitative Motion Analysis software (QMA<sup>®</sup>, Medical Metrics,  
2 Inc., Houston, TX) to place the landmarks. The QMA<sup>®</sup> software has previously been validated <sup>(44-49)</sup>.

3

4 **Figure 1:** Details of anatomic landmark placement. Dashed lines show the assumed mid-sagittal plane of  
5 the superior and inferior endplates, identified as bisecting the radiographic shadows of the left and right  
6 sides of the endplates (yellow arrows). The red circles show the four landmarks used to measure  
7 vertebral body morphology. The red arrow points to an anterior osteophyte that is ignored. The dotted  
8 lines show the anterior and posterior vertebral body heights.



9

10 Details of landmark placement are provided, since as Keynan et al. noted, “There is a lack of  
11 standardization in the literature regarding choice and technique for the measurement of these  
12 parameters.”<sup>(37)</sup>. The standardized landmark placement was focused on the mid-sagittal plane of each  
13 vertebral body (Figure 1), excluding residual uncinat processes, and placing landmarks that identify the  
14 corners of the vertebral bodies prior to any osteophyte formation <sup>(50,51)</sup>. Landmark placement is similar  
15 to that illustrated in Fig 3 of Genant et al <sup>(52)</sup>. The posterior superior landmarks are placed so as to  
16 represent the endplate as it would appear on a mid-sagittal slice of a CT exam, rather than up at the top

1 of the posterior ridges on the upper endplate similar to logic described in Keynan et al. <sup>(37)</sup>. When the X-  
2 ray beam path through a vertebra is not perpendicular to the mid-sagittal plane of the vertebra, the left  
3 and right sides of the vertebral endplates and the left and right aspects of the posterior vertebral body  
4 may be seen on the X-ray.<sup>(16,20)</sup> It is understood that the radiographic shadows that might be interpreted  
5 as the left and right rims of the endplates might not exactly represent the endplate rims<sup>(16)</sup>, but it is  
6 assumed that the line bisecting these shadows is the best-available radiographic approximation of the  
7 mid-sagittal plane. Therefore, the mid-sagittal plane is assumed to be midway between the radiographic  
8 shadows of the left and right sides of endplate, and midway between the radiographic shadows of the  
9 left and right aspects of the posterior wall. Multiple examples of landmark placement are provided  
10 online<sup>3</sup> to help readers appreciate the nuances of how landmarks were placed. Vertebral morphometry  
11 calculated from landmarks is dependent on details of landmark placement, and it is therefore important  
12 to appreciate landmark placement details. The examples provided include vertebrae with abnormal  
13 morphology (as defined by the metrics presented in this paper) and help to appreciate the variety of  
14 conditions that may be identified through utilization of standardized morphology metrics.

15 The neural networks and coded logic produced landmark coordinates for the vertebral bodies  
16 from L1 to S1 in lumbar spine radiographs and C2 to C7 in cervical spine radiographs. An anomaly  
17 detection algorithm was then run on those data to identify anomalies due to very poor image quality,  
18 severe image artifacts, and very unusual anatomy. Those anomalies were removed from subsequent  
19 data analysis.

20

---

<sup>3</sup> <https://www.dropbox.com/sh/qzrocrh86goxarx/AAADRXs5HoEjVdRcGlwA4bela?dl=0>

1           The coordinates of the four vertebral body corners were used to calculate six vertebral body  
2 morphology metrics:

- 3           1) anterior / posterior vertebral body height ratio (VBHR)
- 4           2) superior / inferior endplate width ratio (EPWR)
- 5           3) forward / backward diagonal ratio (FBDR)
- 6           4) height / width ratio (HWR)
- 7           5) angle between the superior and inferior endplates (EPA)
- 8           6) angle between the posterior wall and the superior endplate (PSA)

9  
10           The NHANES-II lumbar radiographs were obtained with the subject in the lateral recumbent  
11 position with body flexion <sup>(53)</sup>. As this is not the typical position used to obtain a radiograph for vertebral  
12 fracture detection, global and local spinal lordosis were not calculated. The NHANES-II cervical spine  
13 radiographs were obtained with the subject standing and a 25 lb weight in each hand to pull down the  
14 shoulders, and the head and neck flexed forward. <sup>(53)</sup>

15           Principal component analysis was used to determine if the dimensionality of the data can be  
16 reduced (Stata Ver 15). The dependency of the six vertebral morphology metrics on sex, age, race,  
17 nation of origin, height, weight, and BMI was assessed using analysis of variance (ANOVA, Stata Ver 15).  
18 In the NHANES-II study, race was recorded as white, black, or other. Nation of Origin was one of 13  
19 categories.

20           Descriptive statistics were tabulated and used to identify and trim outliers in the data that may  
21 compromise the definition of “normal” vertebral morphology <sup>(54-57)</sup>. Outliers  $<> 3$  standard deviations  
22 (SD) from the mean were inspected to identify typical explanations for each type of outlier. The data  
23 were trimmed by excluding all vertebrae where any of the six metrics was  $<> 2$  SD from the mean, and  
24 descriptive statistics were tabulated for use in quantifying deviation from average for each metric.



1 Correlations between the six morphology metrics were also explored. Finally, means and standard  
2 deviations were tabulated for each metric, for each vertebra from L1 to S1 and from C2 to C7. Those  
3 means and standard deviations were then used to calculate standardized metrics (standard deviations  
4 from mean) for the entire data sets. All statistical analysis was completed using Stata ver 15 (College  
5 Station, TX, USA).

6 The NHANES-II study also includes a response to the question: “Have you ever had pain in your  
7 back on most days for at least two weeks”. That data was analyzed for associations with the six  
8 morphology metrics.

9 To assess reproducibility of the metrics, the six morphology metrics were also calculated for  
10 flexion and extension radiographs for 415 asymptomatic volunteers. This is an expanded version of  
11 previously published data.<sup>(58)</sup> The morphology metrics were calculated from landmark coordinates  
12 obtained using the same neural networks and coded logic used with the NHANES-II X-rays. The  
13 morphology of each vertebra should not have changed between flexion and extension, unless there was  
14 an unhealed vertebral fracture, which is unlikely in an asymptomatic person. However, the position of  
15 each vertebrae relative to the central beam path did change substantially between flexion and extension  
16 in most cases. The morphology measured from the flexion X-ray was compared to the morphology  
17 calculated from the extension X-ray using Bland-Altman limits of agreement. Any variability in the  
18 metrics is assumed to be error due to variability in radiographic projection.

19 To further document expected variability in the metrics specifically due to variations in radiographic  
20 projections that can occur in clinical practice, digitally reconstructed radiographs were used. Thin-slice (1  
21 mm or less) anonymized computed tomography exams of the lumbar spine of 26 individuals or cadavers  
22 (from past research studies) were interpolated to 0.2 mm isotropic resolution. The three-dimensional  
23 coordinates of the four anatomic landmarks described in Figure 1 were digitized, in 3D, for the mid-

1 sagittal plane of each vertebra from L1 to S1, using intersecting axial, coronal, and sagittal slices.  
2 Landmarks were thus placed in the mid-sagittal plane at clear and well-defined locations. This may be as  
3 close as possible to “gold standard” landmark placement. The image processing and landmark  
4 digitization was completed using Slicer 3D.<sup>(59)</sup> Custom python code was developed to create 2D, digitally  
5 reconstructed radiographs (DRR) from the 3D data using Plastimatch DRR <sup>(60)</sup>. The projection matrices  
6 used by Plastimatch DRR to create the X-rays were also used to calculate, from the known 3D  
7 coordinates, the precise coordinates of each landmark on the 2D simulated X-ray. A 40 inch source-to-  
8 image distance was used for all X-rays. The base x-ray was centered on the centroid of the L3 vertebra  
9 (this is the baseline isocenter). Forty additional X-rays were generated with the following variations:

- 10 • random beam tilts of between  $\pm 10$  deg in each of the sagittal and axial planes
- 11 • random displacements from the baseline isocenter  $\pm 2$  endplate widths (EPW) in the AP and LR  
12 directions, plus  $\pm 3$  EPW shifts in the cranial-caudal direction
- 13 • random image rotations between  $\pm 24$  deg

14 These variations resulted in a wide range of radiographic projections. The landmarks in every image  
15 were precisely known so there were zero errors in landmark coordinates relative to the 3D coordinates.  
16 This allows for documenting the effect of variability in vertebral morphology metrics that are due to  
17 radiographic projection, without uncertainty in landmark placement. These precisely calculated  
18 landmarks were analyzed to obtain the six vertebral body morphology metrics. There were thus 41  
19 morphology measurements for every vertebra from L1 to S1 for each of 26 independent CT exams. The  
20 variability in each metric was assessed using the standardized version of the metric, since all of these  
21 metrics are thereby expressed as standard deviations from the average and that allows for combining  
22 data for all vertebrae and for consistent interpretation of the data. Although the coordinates of  
23 landmarks in every simulated X-ray were precisely calculated, the quality of the simulated radiographs

1 was below clinical standards and not appropriate for use toward validating the accuracy of neural  
2 networks in obtaining landmarks.

### 3 Results and Discussion

4 Due to the large quantity of data to be presented, the results and discussion are  
5 combined.

#### 6 Summary of Data Analyzed

7 Computer generated lumbar landmarks were obtained for 42,980 vertebrae from 7,364 of 7,423  
8 total lumbar radiographs and 54,093 vertebrae from 9,662 out 9,669 cervical radiographs for subjects 25  
9 to 74 years old. 50 lumbar X-rays were not analyzed due to an error in file transfer and seven cervical  
10 and nine lumbar X-rays were non-analyzable. Landmark generation for all lumbar and cervical  
11 radiographs took 22 hrs on a dual CPU, 4 GPU (Nvidia T4) production server. The age range was 25 to 74  
12 years and the demographic and body size data are summarized in Table 1.

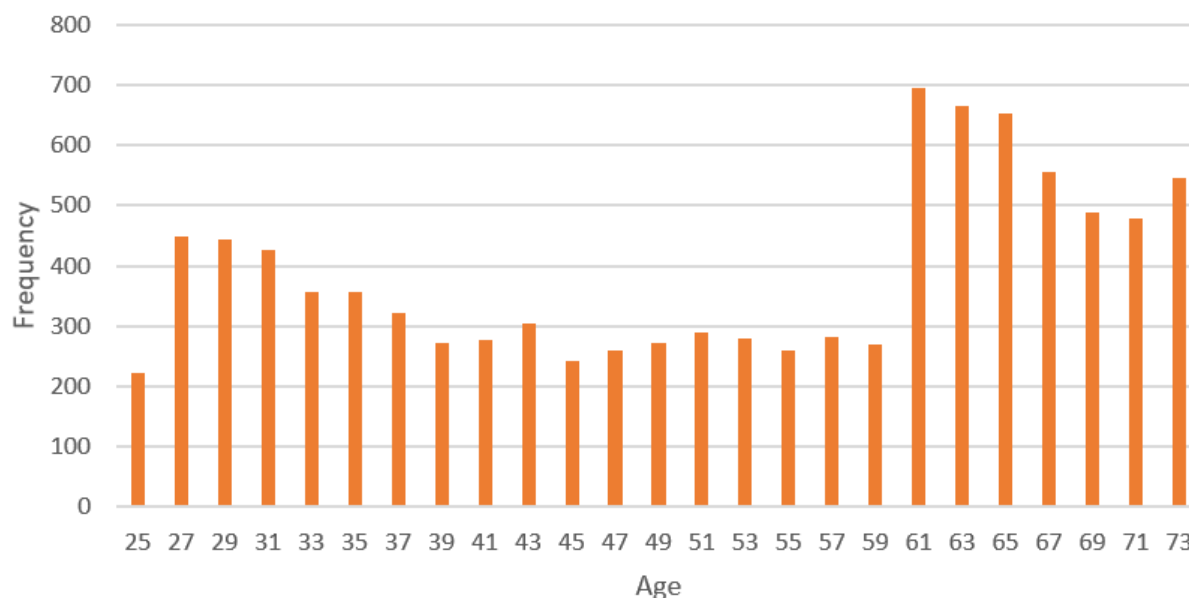
13 Table 1: Sample size, average age and [SD], and average BMI (kg/m<sup>2</sup>) and [SD]. Age and BMI were  
14 missing for two subjects.

	Lumbar			Cervical		
Sex	N	Age	BMI	N	Age	BMI
Male	4,553	50.9[15.3]	25.5[4.0]	4,610	50.8[15.3]	25.5[4.0]
Female	2,809	63.3[6.4]	26.4[5.5]	5,050	51.4[15.2]	25.8[5.8]

15  
16 Note that females are under-represented in the lumbar data. By design of the NHANES-II study,  
17 no lumbar X-rays were taken for pregnant women or women under 50.<sup>(61)</sup> An additional limitation

1 relates to the data on race and nation of origin recorded in the NHANES-II data. Based on the cervical  
2 data, 86.9% of NHANES-II subjects were “white”, 11.2% “black” and the rest “other”. A more uniform  
3 representation of races would likely be needed to fully understand the importance of race. The same is  
4 true for the “nation of origin” data in NHANES-II. NOTE: the language in quotes is as used in the  
5 NHANES-II documentation: 74.1% of subjects are classified as “OTHER EUROPE SUCH AS GERMAN,  
6 FRENCH, ENGLISH, IRISH”, 10.8% classified as “BLACK, NEGRO OR AFRO-AMERICAN”, and the rest  
7 distributed in 1 of 11 other classifications. Djoumessi et al. have reported differences in the VBHR  
8 between different nations of origin <sup>(56)</sup>. It would be valuable to repeat this analysis on X-rays that better  
9 represent races and nations of origin not well represented in the NHANES-II data. The ages of the  
10 subjects (using cervical data) were also biased toward older ages (Figure 5). That may explain the  
11 relatively high rate of vertebrae with at least one abnormal morphology metric.

12 Figure 5: Distribution of ages in the NHANES-II study.



13  
14 Principal component analysis (PCA) of the lumbar data documented that 5 primary components  
15 can explain 99.8% of the variance in the morphology data. PCA of the cervical data documents that 6

1 components explain 99.7% of the variance. Since PCA does not appreciably reduce the dimensionality of  
2 the data, it was not investigated further.

### 3 Effect of covariates

4 Analyses of variance (after excluding outliers) revealed that age, sex, race, and nation of origin  
5 were statistically significant for most of the variables (Tables 2 and 3), and the  $R^2$  was high. However, the  
6 F-statistic was generally small for all variables except vertebral level. Excluding all variables except  
7 vertebra changed the  $R^2$  very little (e.g. for the lumbar data, VBHR  $R^2$  went from 0.848 to 0.842 and HWR  
8  $R^2$  went from 0.839 to 0.800). This documents that age, sex, race and body size are significant, but  
9 contribute relatively little to the variability in the morphology metrics. Therefore, except for vertebral  
10 level, all data were pooled for subsequent analysis. Previous studies have shown a dependence of  
11 vertebral morphology on age, gender, race and other variables, but the relative magnitude of the  
12 influence was not described.<sup>(62,63)</sup>

### 13 Descriptive Statistics for Normative Reference

14 Descriptive statistics for the morphologic parameters are provided in Tables 4 and 5. Mean  
15 values changed very little for the morphology metrics after removing the outliers (< 2 SD from mean),  
16 suggesting that normal morphology is dominant in the data. Standard deviations were reduced after  
17 removing outliers. These observations are the same as reported in a similar study<sup>(56)</sup>. Data with a  
18 Gaussian distribution have a skewness of 0 and kurtosis of 3. Excluding outliers substantially reduced the  
19 skewness from an average of 0.27 to 0.05, and substantially narrowed the tails of the distribution from  
20 an average kurtosis of 4.1 to 2.5. This documents that the distribution of data was more Gaussian after  
21 removing outliers. There was remarkably small variation in vertebral morphology after accounting for  
22 vertebra (L1 , L2, etc.). The PSA, FBDR, and EPWR generally had the lowest coefficients of variation  
23 (Tables 6 and 7).

1 **Table 2:** Results of multivariate analysis of variance to determine the importance of multiple independent  
 2 variables for explaining variability in the lumbar vertebral morphology metrics. The “F” columns provide the “F”  
 3 statistic. The “P>F” columns provide the statistical significance. The F values document how much of the variability  
 4 in the dependent variable is explained by each independent variable.

Variable	VBHR R <sup>2</sup> = 0.85		EPWR R <sup>2</sup> = 0.92		FBDR R <sup>2</sup> = 0.93		HWR R <sup>2</sup> = 0.84		EPA R <sup>2</sup> = 0.82		PSA R <sup>2</sup> = 0.75	
	F	P > F	F	P > F	F	P > F	F	P > F	F	P > F	F	P > F
<i>Vertebra</i>	39009	<b>0.000</b>	85907	<b>0.000</b>	94108	<b>0.000</b>	34955	<b>0.000</b>	32422	<b>0.000</b>	20937	<b>0.000</b>
Age	54	<b>0.000</b>	133	<b>0.000</b>	72	<b>0.000</b>	1414	<b>0.000</b>	8	<b>0.004</b>	10	<b>0.002</b>
Sex	350	<b>0.000</b>	82	<b>0.000</b>	4	<b>0.054</b>	3969	<b>0.000</b>	535	<b>0.000</b>	146	<b>0.000</b>
Race	7	<b>0.001</b>	0	0.716	2	0.218	6	<b>0.002</b>	8	<b>0.000</b>	4	<b>0.021</b>
Nation of Origin	5	<b>0.000</b>	1	0.398	2	<b>0.034</b>	13	<b>0.000</b>	6	<b>0.000</b>	4	<b>0.000</b>
Weight	0	0.995	3	0.089	0	0.857	10	<b>0.002</b>	0	0.760	0	0.506
Height	0	0.829	7	<b>0.010</b>	2	0.211	2	0.194	1	0.458	0	0.526
BMI	0	0.645	3	0.076	0	0.868	0	0.482	0	0.911	0	0.496

5

6 **Table 3:** Results of multivariate analysis of variance to determine the importance of multiple independent  
 7 variables for explaining variability in the cervical vertebral morphology metrics.

Variable	VBHR R <sup>2</sup> = 0.72		EPWR R <sup>2</sup> = 0.77		FBDR R <sup>2</sup> = 0.68		HWR R <sup>2</sup> = 0.88		EPA R <sup>2</sup> = 0.72		PSA R <sup>2</sup> = 0.55	
	F	P > F	F	P > F	F	P > F	F	P > F	F	P > F	F	P > F
<i>Vertebra</i>	22659	<b>0.000</b>	30247	<b>0.000</b>	19163	<b>0.000</b>	66488	<b>0.000</b>	22852	<b>0.000</b>	10977	<b>0.000</b>
Age	136	<b>0.000</b>	161	<b>0.000</b>	421	<b>0.000</b>	2683	<b>0.000</b>	103	<b>0.000</b>	415	<b>0.000</b>
Sex	0	0.553	14	<b>0.000</b>	381	<b>0.000</b>	829	<b>0.000</b>	3	0.065	236	<b>0.000</b>
Race	7	<b>0.001</b>	0	0.928	0	0.895	27	<b>0.000</b>	7	<b>0.001</b>	1	0.472
Nation of Origin	2	<b>0.017</b>	1	0.223	7	<b>0.000</b>	21	<b>0.000</b>	2	<b>0.039</b>	6	<b>0.000</b>
Weight	18	<b>0.000</b>	5	<b>0.029</b>	0	0.958	4	<b>0.044</b>	24	<b>0.000</b>	1	0.303
Height	27	<b>0.000</b>	6	<b>0.017</b>	25	<b>0.000</b>	51	<b>0.000</b>	34	<b>0.000</b>	32	<b>0.000</b>
BMI	14	<b>0.000</b>	5	<b>0.033</b>	0	0.752	0	0.528	18	<b>0.000</b>	0	0.606

- 1 Table 4: Means and [standard deviations] for 35,275 lumbar vertebrae. All vertebrae where any metric  
 2 was  $\leq$  2 Std Dev from the average for all vertebrae in the NHANES-II study were excluded.

Vert	VBHR	EPWR	FBDR	HWR	EPA	PSA
L1	0.944[0.043]	0.994[0.017]	0.994[0.017]	0.823[0.059]	-2.678[2.035]	88.499[1.250]
L2	0.979[0.040]	0.999[0.018]	0.973[0.019]	0.827[0.059]	-1.021[1.907]	87.919[1.251]
L3	0.997[0.039]	1.009[0.017]	0.960[0.016]	0.821[0.058]	-0.126[1.825]	87.222[1.266]
L4	1.032[0.046]	1.003[0.019]	0.938[0.017]	0.799[0.058]	1.406[2.049]	86.753[1.383]
L5	1.154[0.066]	1.023[0.025]	0.930[0.023]	0.788[0.058]	6.517[2.655]	87.671[1.970]
S1	1.296[0.071]	1.611[0.154]	1.268[0.067]	1.157[0.086]	11.937[3.456]	96.564[3.539]

- 3  
 4 Table 5: Means and [standard deviations] for 44938 cervical vertebrae. All vertebrae where any metric  
 5 was  $\leq$  2 Std Dev from the average for all vertebrae in the NHANES-II study were excluded.

Vertebra	VBHR	EPWR	FBDR	HWR	EPA	PSA
C2	1.130[0.042]	0.790[0.047]	0.799[0.027]	1.324[0.083]	6.389[2.833]	85.069[2.484]
C3	0.965[0.047]	0.935[0.034]	0.802[0.029]	0.899[0.076]	-2.810[2.484]	78.307[2.518]
C4	0.949[0.045]	0.945[0.032]	0.825[0.027]	0.855[0.069]	-3.322[2.265]	79.320[2.375]
C5	0.927[0.043]	0.929[0.033]	0.848[0.029]	0.819[0.070]	-4.354[2.101]	81.006[2.460]
C6	0.955[0.044]	0.952[0.030]	0.879[0.028]	0.792[0.072]	-2.521[2.060]	83.011[2.250]
C7	0.998[0.035]	1.009[0.030]	0.921[0.025]	0.856[0.064]	-0.108[1.730]	84.868[1.811]

6

- 1 Table 6: Coefficients of variation (CV = SD / mean) for 35,275 lumbar vertebrae. Note that the large
- 2 values for L3 EPA is due to the mean value being very close to zero (ref Table 4) and the CV is difficult to
- 3 interpret when the denominator is near zero.

Vertebra	VBHR	EPWR	FBDR	HWR	EPA	PSA	N
L1	0.045	0.018	0.017	0.072	-0.760	0.014	5,870
L2	0.041	0.018	0.019	0.072	-1.868	0.014	5,987
L3	0.039	0.017	0.017	0.070	-14.484	0.015	5,946
L4	0.045	0.019	0.018	0.072	1.458	0.016	5,953
L5	0.057	0.024	0.025	0.074	0.407	0.022	5,886
S1	0.055	0.095	0.053	0.074	0.290	0.037	5,638

4

- 5 Table 7: Coefficients of variation (SD / mean) for 44,938 cervical vertebrae.

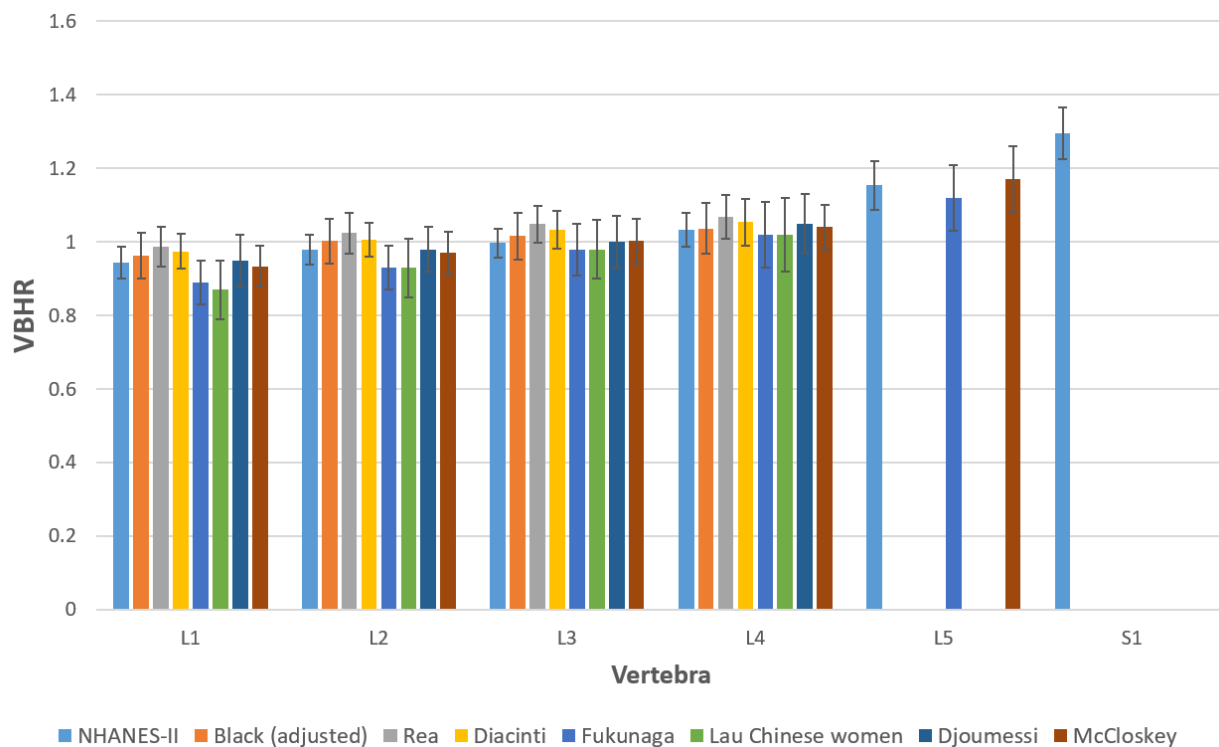
Vertebra	VBHR	EPWR	FBDR	HWR	EPA	PSA	N
C2	0.037	0.059	0.034	0.063	0.443	0.029	7,705
C3	0.049	0.036	0.036	0.085	-0.884	0.032	8,180
C4	0.047	0.034	0.033	0.081	-0.682	0.03	8,017
C5	0.046	0.036	0.035	0.086	-0.482	0.03	8,127
C6	0.046	0.032	0.032	0.091	-0.817	0.027	7,695
C7	0.035	0.03	0.027	0.074	-16.01	0.021	5,214

6

- 7 Figure 6 provides comparison of VBHR data in Table 4 to some previously reported morphology
- 8 metrics.<sup>(54-57,64-66)</sup>



1 Figure 6: Comparison of previously published<sup>(54,55,57,64-66)</sup> VBHR data to data obtained from the NHANES-  
 2 II study.



3  
 4 The six morphologic variables were significantly correlated with each other (Tables 8 and 9). EPA  
 5 and VBHR were almost perfectly correlated, suggesting that only one of these two variables needs to be  
 6 analyzed. Although VBHR and EPA were almost perfectly correlated, the other metrics were not. That  
 7 observation supports that the metrics each independently describe some unique aspect of vertebral  
 8 morphology. It cannot be determined from the NHANES data alone which of the six metrics will prove  
 9 most valuable in research and clinical practice. It may be worth exploring the utility of each metric, and  
 10 combinations of metrics, in future studies where the goal is to diagnose fractures or deformities, or to  
 11 predict response to treatment.

1 Table 8: Pearson’s correlation coefficients between morphologic variables for 35,275 lumbar vertebrae,  
 2 after excluding vertebrae where any of the morphologic variables was  $\pm 2$  SD from the average for  
 3 42,980 vertebrae.

	VBHR	EPWR	FBDR	HWR	EPA	PSA
VBHR	1					
EPWR	0.7741	1				
FBDR	0.6694	0.9373	1			
HWR	0.6661	0.8844	0.8712	1		
EPA	0.9904	0.7338	0.6205	0.6551	1	
PSA	0.7394	0.7911	0.9069	0.8069	0.7188	1

4  
 5 Table 9: Pearson’s correlation coefficients between morphologic variables for 44,938 cervical vertebrae,  
 6 after excluding vertebrae where any of the morphologic variables was  $\pm 2$  SD from the average for  
 7 54,093 vertebrae.

	VBHR	EPWR	FBDR	HWR	EPA	PSA
VBHR	1					
EPWR	-0.6536	1				
FBDR	-0.2641	0.5298	1			
HWR	0.7861	-0.7548	-0.4464	1		
EPA	0.9848	-0.5813	-0.1828	0.7646	1	
PSA	0.5704	-0.352	0.5461	0.3203	0.6047	1

8  
 18

1 Multiple prior publications report efforts to establish reference data that can be used to help  
2 diagnose vertebral body fractures. <sup>(6-10,12-15,17-43)</sup> Although multiple metrics need to be considered, such  
3 as intervertebral angles, listhesis and canal occlusion <sup>(37)</sup>, compression of the vertebral body height is  
4 commonly used to identify fractures. When severe loss of vertebral body height has occurred, clinicians  
5 can unanimously agree that a fracture has occurred, but when the fracture is subtle, reference data may  
6 be important. Mid-vertebral landmarks that could be used to measure mid-vertebral height were not  
7 routinely available in the images used to train the neural networks, so mid-vertebral body heights could  
8 not be obtained automatically. Change in vertebral shape is considered one of the best indicators of a  
9 fracture <sup>(6)</sup>. The NHANES-II images provide only one radiograph per individual, so change in shape cannot  
10 be explored. Nevertheless, the low variability of vertebral shape within a large population supports the  
11 potential for using a single X-ray to detect abnormal morphology with reference to the NHANES-II data.  
12 It is also likely, that with advancements in the application of artificial intelligence to producing spine  
13 metrics, the data in Tables 4 and 5 may provide valuable quality control data to determine if the AI  
14 predictions are aberrant or if caution is needed in interpreting the results.

15 Application of the vertebral morphology metrics, such as in Tables 4 and 5, to the challenge of  
16 documenting vertebral fractures will require defining a threshold level of these metrics that can be used  
17 to document the presence/absence of a fracture. Eastell et al. and McCloskey et al. used a threshold of 3  
18 SD for the VBHR <sup>(64,67)</sup>. Black et al. used a threshold of 20% difference from normal VBHR <sup>(25)</sup>. Genant et  
19 al. used the criteria of a 15% difference compared to the mean VBHR of a normal population <sup>(52)</sup>. The  
20 average VBHR in Table 4 for L1 to L4 is about 1 and one SD is about 0.042, so a 15% difference as used  
21 by Genant et al. would be equivalent to approximately a VBHR Z-score of 3.5. A gold-standard test for a  
22 true vertebral fracture does not exist, so validating a threshold remains a challenge. An alternative,  
23 treatment specific approach, would be to validate a threshold that predicts a positive response to  
24 treatment. That approach could potentially be retrospectively tested at low cost using automated

1 measurements of morphology applied to imaging and data from prior studies that collected appropriate  
 2 imaging.

### 3 Differences between levels

4 It is reasonable to hypothesize that differences in morphology between adjacent levels may help  
 5 to identify abnormal morphology. For example, if one level has an anterior wedge fracture, then a  
 6 significant difference in the VBHR would be expected when compared to the immediately adjacent  
 7 levels. A problem occurs when adjacent vertebrae are both fractured or otherwise abnormal.  
 8 Comparison to adjacent levels has been a criteria used or proposed by many investigators, so data that  
 9 could be used to implement adjacent vertebrae comparisons are provided in tables 10 and 11. For  
 10 example, using data in Table 10, if the VBHR is to be compared between L4 and L5 vertebrae, and  
 11 normal variability is defined as anything within the 95% confidence interval for data from the NHANES-II  
 12 study, than the VBHR when comparing L4 to L5 should not vary by more than  $0.124 \pm 1.96 * 0.059$ , or  
 13 0.007 to 0.24.

14 **Table 10:** The means and [standard deviations] for the absolute differences between adjacent lumbar  
 15 vertebrae for each of the vertebral morphology metrics. These data include only those vertebrae with  
 16 no abnormalities in any morphology metric using the data in Table 4. The sample size is smaller, since if  
 17 either vertebra had any abnormality, that comparison between levels was excluded.

Compare	N	VBHR	EPWR	FBDR	HWR	EPA	PSA
L1 to L2	3540	0.041[0.030]	0.017[0.013]	0.024[0.016]	0.023[0.018]	2.009[1.442]	1.127[0.864]
L2 to L3	3635	0.034[0.026]	0.018[0.014]	0.018[0.013]	0.023[0.019]	1.611[1.224]	1.136[0.847]
L3 to L4	3628	0.043[0.031]	0.018[0.013]	0.025[0.015]	0.028[0.021]	1.955[1.410]	1.106[0.838]
L4 to L5	3554	0.124[0.059]	0.026[0.019]	0.020[0.015]	0.027[0.022]	5.191[2.443]	1.583[1.178]
L5 to S1	3137	0.145[0.080]	0.589[0.134]	0.342[0.061]	0.372[0.074]	5.733[3.542]	9.113[3.504]

18

1 Table 11: The means and [standard deviations] for the absolute differences between adjacent cervical  
2 vertebrae for each of the vertebral morphology metrics. These data include only those vertebrae with  
3 no abnormalities in any morphology metric using the data in Table 5.

Compare	N	VBHR	EPWR	FBDR	HWR	EPA	PSA
C2 to C3	4636	0.167[0.053]	0.145[0.053]	0.025[0.018]	0.426[0.070]	9.297[3.241]	6.736[2.868]
C3 to C4	5039	0.041[0.031]	0.032[0.024]	0.027[0.019]	0.052[0.038]	2.063[1.553]	1.983[1.475]
C4 to C5	5023	0.041[0.031]	0.033[0.025]	0.028[0.020]	0.047[0.034]	2.046[1.509]	2.207[1.607]
C5 to C6	4850	0.045[0.033]	0.036[0.026]	0.034[0.022]	0.044[0.034]	2.429[1.732]	2.466[1.719]
C6C7	3231	0.051[0.035]	0.060[0.036]	0.045[0.025]	0.064[0.040]	2.713[1.768]	2.363[1.619]

4

## 5 Abnormalities in the NHANES data

6 Mean and standard deviation data from Table 4 and Table 5 were used to calculate the vertebra  
7 specific Z-score for each metric according to the following equation

$$8 \quad Z = \frac{x - \mu}{\sigma}$$

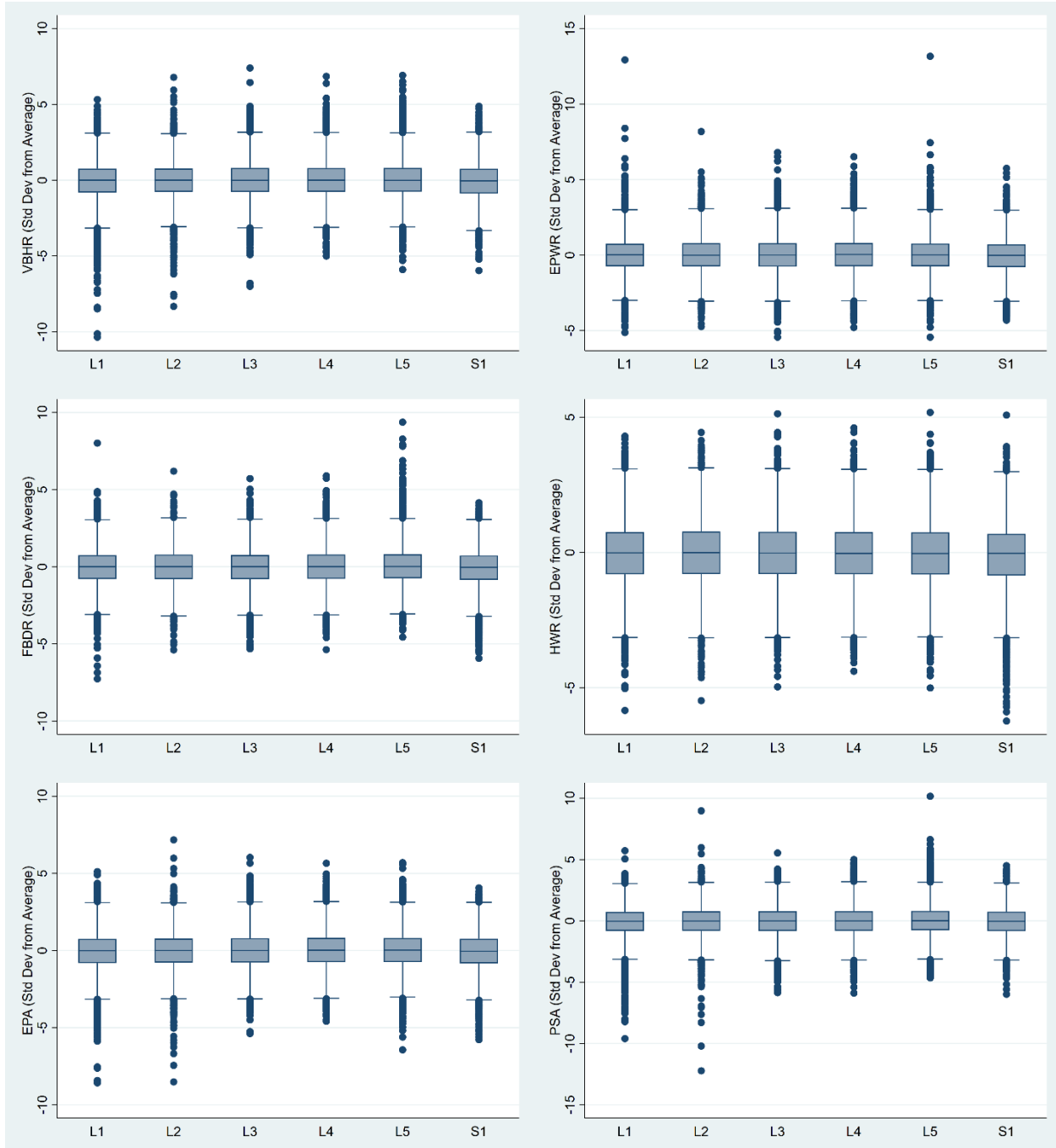
9 where  $x$  is the metric for a vertebra,  $\mu$  is the vertebra specific mean for the corresponding metric, and  $\sigma$   
10 is the vertebra specific standard deviation for the corresponding metric. Based on these Z-scores, 28,813  
11 of 42,980 lumbar vertebrae (67%) were within  $\pm 2$  SD of average (Table 4) for all six of the metrics. These  
12 number are less than obtained after trimming since the standard deviations in Table 4 are smaller than  
13 originally used to identify outliers, and are therefore more sensitive to abnormalities. Similarly, 36,869  
14 of 54,093 cervical vertebrae (68%) were within  $\pm 2$  SD of average (Table 5) for all six of the metrics.

15 Box and whisker plots illustrate the magnitude of outliers that exist in the NHANES data (Figures  
16 7 and 8). Note that the medians are generally centered within the interquartile range and the whiskers

1 are symmetric. Positive VBHR outliers represent posterior wedge-shaped vertebral bodies, while  
2 negative VBHR outliers represent anterior wedge-shaped vertebrae. The wedging may either be from  
3 fractures, remodeling, or congenital conditions and identifying a specific cause would require more than  
4 a single radiograph. Positive EPWR outliers represent vertebrae that are unusually wide superiorly  
5 versus inferiorly, and negative outliers are the opposite. Positive and negative FBDR outliers generally  
6 appear to be best explained by remodeling, perhaps after a fracture that caused splaying of the fracture  
7 fragments. Positive HWR outliers appear as unusually tall vertebrae, while negative HWR outliers  
8 generally appear to be best explained by a crush fracture. Positive EPA outliers appear as anteriorly  
9 wedge-shaped vertebrae while negative outliers appear as posteriorly wedge-shaped vertebrae. Positive  
10 PSA outliers generally appear as posterior wedge-shaped vertebrae and negative outliers generally  
11 appear as anteriorly wedge-shaped vertebrae. It is not possible to definitively identify the cause of each  
12 abnormality from a single radiograph. Examples of each of high positive and low negative outliers, for  
13 each of the six metrics, including landmarks, can be viewed online.  
14 (<https://www.dropbox.com/sh/qzrocrh86goxarx/AAADRxs5HoEjVdRcGIwA4bela?dl=0>).

15

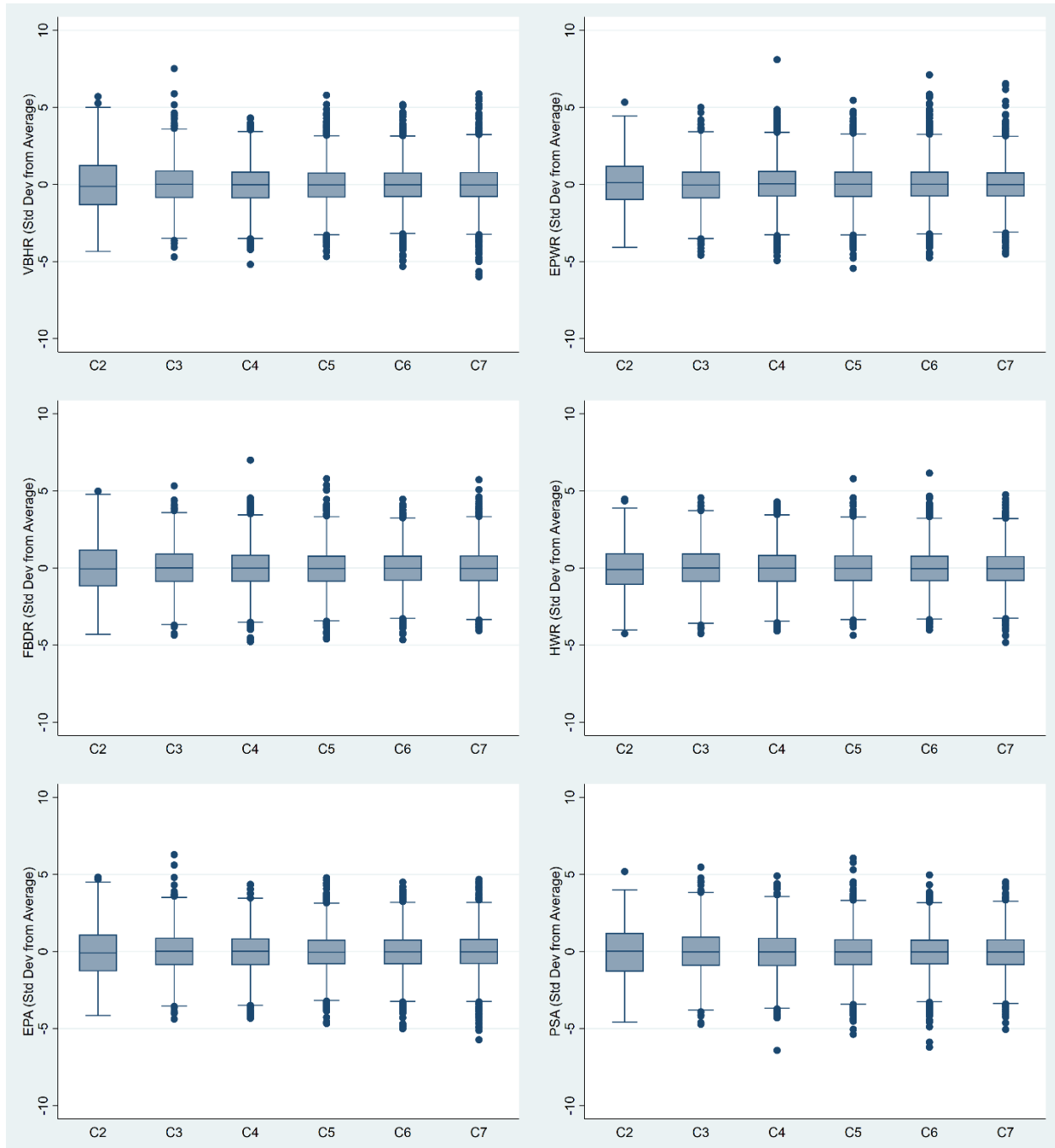
- 1 Figure 7: Box and whisker plots for the six morphology metrics calculated for 42,980 lumbar vertebrae.
- 2 These plots document the types and magnitude of outliers in the data.



3

4

- 1 Figure 8: Box and whisker plots for the six morphology metrics calculated for 54,093 cervical vertebrae.
- 2 These plots document the types and magnitude of outliers in the data.



3

4



## 1 Error analysis for the morphology metrics

2 One advantage of the six vertebral morphology metrics is the lack of dependence on  
3 radiographic magnification. All metrics are ratios of single vertebra dimensions or angles between  
4 anatomic reference lines and should not vary appreciably with radiographic magnification (assuming  
5 sufficient spatial resolution). This is important since radiographic magnification can vary substantially  
6 between lumbar X-rays<sup>(78-80)</sup> and this can result in errors in measurements of vertebral dimensions<sup>(26)</sup>.

7 Some error in the data from Tables 4 and 5 may occur as a result of the challenges of vertebral  
8 labeling. A systematic approach to definitively labeling vertebrae requires whole spine imaging<sup>(68)</sup>. An X-  
9 ray or CT of the entire spine from occiput to sacrum was not available (X-rays typically included T10 to  
10 S1), so it cannot be certain that vertebrae were optimally labeled if there were <> 5 lumbar vertebrae  
11 transitional vertebrae, or other abnormalities<sup>(69,70)</sup>. Spines with <> 5 lumbar vertebrae are  
12 uncommon<sup>(71)</sup>. Transitional vertebrae may be more common<sup>(72)</sup>, but it is unclear whether this would  
13 consistently compromise labeling from a lateral spine x-ray. The assumption is that the vertebrae were  
14 correctly labeled in the majority of cases and the possibility of mislabeled vertebrae in a minority of  
15 cases would not significantly affect the results. The prevalence of transitional vertebrae reported in the  
16 peer-reviewed literature varies between 4 and 36%<sup>(73-77)</sup>. The prevalence in the NHANES-II radiographs  
17 is unknown.

18 The six morphology metrics calculated from flexion X-rays were compared to the metrics  
19 calculated from paired extension X-rays for 2,132 vertebrae from L1 to S1. There should be no difference  
20 in vertebral morphology between flexion and extension (assuming no mobile vertebral body fractures).  
21 Table 12 provides the results of the Bland-Altman limits of agreement analysis. This variability is  
22 assumed due to the differences in radiographic projection between images. Typically, there would not  
23 be much difference in radiographic projection between carefully obtained flexion versus extension X-

1 rays, and this is reflected in the mean differences in Table 12, but substantial differences did occur in  
2 some subjects, and this is reflected in the low and high limits of agreement in Table 12. This experiment  
3 documents that some morphology metrics could vary by as much as 1.65 Std Dev due to radiographic  
4 projection differences. Thus, it is important to strive for radiographs where the vertebrae of interest are  
5 positioned near the center of the radiograph. The experiment with the simulated X-rays, discussed  
6 below, compliments and helps to explain results in Table 12.

7 [Table 12](#): Results of a Bland-Altman limits of agreement analysis for morphology (expressed as SD from  
8 the average) measured from 2,132 lumbar flexion X-rays compared to paired extension X-rays. The  
9 mean differences between the two measurements, as well as the low and high limits of agreement are  
10 provided. These data are specific to the methods used to obtain anatomic landmarks from the  
11 radiographs.

Metric	Mean Difference	Low	High
VBHR	0.15	-0.93	1.24
EPWR	0.028	-1.59	1.65
FBDR	-0.09	-1.58	1.4
HWR	-0.016	-0.92	0.89
EPA	0.16	-0.99	1.3
PSA	0.02	-1.57	1.61

12

### 13 [Variability due to radiographic projection](#)

14 The experiment using simulated X-rays was completed to help understand how variability in  
15 radiographic projection can affect the vertebral morphology metrics. This could be important, for  
16 example, with sub-optimally positioned patients, or in patients with scoliosis or other deformities, all of

1 which can result in the x-ray beam passing from source to image obliquely to the vertebral endplates or  
2 the posterior wall of the vertebrae. There should ideally be no variability in vertebral morphology in the  
3 experiment using the simulated x-rays, since all x-rays (for a subject) were created from the same CT  
4 exam. All landmarks were precisely calculated from one set of 3D landmark coordinates, based on the  
5 geometry of the radiographic projection. Any variability is due to the effect of variability in radiographic  
6 projection on calculation of morphology. Variability was assessed using the standardized version of the  
7 metrics so that all metrics have the same units (standard deviations from average) and can thus be more  
8 easily compared. Conventional measures of variability, such as the coefficient of variation ( $CV = \text{std dev} /$   
9  $\text{mean}$ ) or the coefficient of quartile variation ( $CQV = 100 * (Q_3 - Q_1) / (Q_3 + Q_1)$ ) were not used as the  
10 mean values and  $(Q_3 + Q_1)$  were close to zero for some of the standardized metrics and CV or CQV are  
11 uninterpretable when the denominator approaches zero. Instead, the median and the range (max –  
12 min) for the standardized metrics was calculated for the 41 simulated X-rays that were generated from  
13 each CT exam. The median and 95<sup>th</sup> percentile for these ranges, across the 26 CT exams, are provided in  
14 Table 13. In Table 13, the observed ranges of standardized metrics due to variability in radiographic  
15 projection was largest for FBDR and PSA, although the experiment demonstrates that all of the  
16 morphology metrics can vary by over 1 std dev, solely due to variability in radiographic projection. This is  
17 not as much a weakness in the metrics as it is evidence to use caution when interpreting these metrics  
18 in the presence of substantial out-of-plane imaging. It is of course possible that strategies can be  
19 developed for mitigating the effect of variability in radiographic projection. One option would be a  
20 neural network that determines the type and magnitude of out-of-plane and uses that data to correct  
21 the morphology metrics. It is assumed that the error sources discussed above were random sources of  
22 error and would not appreciably affect the means and standard deviations in Table 4 and Table 5.

23

24

1 Table 13: Variability in standardized metrics (Std Dev from Average Normal) due to variability in  
2 radiographic projection. Each of the morphology metrics was calculated for each of the 41 radiographic  
3 projections generated from each CT exam. The range in each metric (5<sup>th</sup> to 95<sup>th</sup> percentile) was then  
4 calculated over all 41 random variations in radiographic projection. This table has the median and 95<sup>th</sup>  
5 percentile variability over the 26 CT exams. For example, due to a wide range in radiographic  
6 projections, EPA will typically vary by 0.37 Std Devs but could vary by as much as 1.1 Std Devs from the  
7 average EPA in a normal vertebra.

Metric	Median	95 <sup>th</sup> Percentile
EPA	0.3658	1.0796
EPWR	0.6643	2.2362
FBDR	2.4421	4.1485
HWR	0.6378	1.1342
PSA	2.1198	3.7636
VBHR	0.3628	1.2465

8

## 9 Additional observations

10 Based on logistic regression, the subject response to the question: “Have you ever had pain in  
11 your back on most days for at least two weeks” was significantly ( $P < 0.05$ ) associated with age, race, and  
12 BMI. Since there is no way to determine which vertebra (if any) might be causing symptoms, the sum of  
13 abnormalities for each subject (metrics  $< -2$  or  $> 2$  SD from the average based on Table 4) from L1 to S1  
14 was calculated. This sum of abnormalities for the EPWR and the FBDR was significantly associated with  
15 the response to the “ever had back pain” question ( $P < 0.05$ ), although the  $R^2$  for these regressions was  
16 very small ( $> 0.01$ ), indicating a very weak association.

1            Of the six morphology metrics, FBDR was the most sensitive and specific metric (area under ROC  
 2 curve=0.995) to use for the purpose of using landmarks to classify vertebrae as S1 vs L1-L5. A threshold  
 3 level of FBDR>1.0574 misclassified only 278/42980 vertebrae as S1 vs L1-L5. Including all data for all  
 4 vertebrae, the proportions of vertebrae with abnormal morphology (< > 2 SD from average based on  
 5 Tables 4 and 5) are summarized in Tables 14 and 15.

6 **Table 14:** Percent of lumbar vertebrae in the NHANES-II study where the morphology metric was either  
 7 > 2 or < -2 SD from the average, using data from Table 4 for the average and SD, and including all  
 8 vertebrae.

Vert	Freq	VBHR		EPWR		FBDR		HWR		EPA		PSA	
		< -2	> 2	< -2	> 2	< -2	> 2	< -2	> 2	< -2	> 2	< -2	> 2
L1	7,037	7.1	3.3	3.9	4.9	4.8	4.0	4.9	4.0	6.5	3.5	6.8	3.6
L2	7,324	5.5	4.8	3.9	5.0	4.9	4.2	4.6	4.3	5.1	4.9	5.5	4.3
L3	7,318	4.5	5.8	3.9	5.1	5.2	4.2	4.5	4.3	4.2	5.6	5.4	4.8
L4	7,333	3.6	5.8	3.8	5.4	4.5	4.4	4.7	4.0	3.9	5.4	5.0	5.0
L5	7,181	3.9	5.9	3.5	5.4	3.2	6.1	4.8	4.1	4.4	5.4	4.4	6.4
S1	6,787	6.1	4.6	8.7	3.8	8.7	3.8	8.1	3.0	5.5	3.9	5.5	4.0
All	42,980	5.1	5.0	4.6	4.9	5.2	4.5	5.2	4.0	4.9	4.8	5.4	4.7

9

10

1 Table 15: Percent of lumbar vertebrae in the NHANES-II study where the morphology metric was either  
 2 > 2 or < -2 SD from the average, using data from Table 5 for the average and SD, and including all  
 3 vertebrae.

Vert	Freq	VBHR		EPWR		FBDR		HWR		EPA		PSA	
		< -2	> 2	< -2	> 2	< -2	> 2	< -2	> 2	< -2	> 2	< -2	> 2
C2	9,647	4.9	5.3	3.8	5.1	4.3	5.3	4.4	4.1	4.9	4.8	5.3	4.3
C3	9,648	3.8	4.5	4.9	3.7	3.8	4.2	4.0	4.2	3.7	4.1	3.7	4.5
C4	9,654	4.6	4.3	4.2	4.5	3.9	4.5	4.4	3.9	4.6	3.9	4.7	4.2
C5	9,613	4.2	3.8	4.4	4.5	4.3	3.8	4.1	3.8	4.1	3.7	4.8	3.4
C6	9,148	4.5	4.0	3.7	4.9	4.1	4.2	4.4	4.0	4.2	3.9	4.6	3.3
C7	6,383	5.0	4.6	3.9	4.9	4.6	4.1	4.6	4.3	4.8	4.3	5.7	4.4
All	54,093	4.5	4.4	4.1	4.6	4.1	4.4	4.3	4.0	4.4	4.1	4.8	4.0

4

## 5 Conclusion

6 The lateral spine lumbar and cervical radiographs from the NHANES-II study were used to  
 7 establish normative vertebral morphology reference data. These data can be used to help objectively  
 8 identify vertebral body fractures, with the assumption that fractures will result in significant deviations  
 9 from large population-based average vertebral body dimensions. Early diagnosis of fracture allows for  
 10 early evaluation and prevention ultimately decreasing deformity as well as risk of future fracture. The  
 11 data in Table 4 and Table 5 can be used to produce standardized metrics in units of standard deviation  
 12 from average. This allows for easy identification of abnormalities resulting from vertebral fractures,  
 13 atypical vertebral body morphologies, and other congenital or degenerative conditions, and removes  
 14 the effect of vertebra allowing data for all vertebrae to be pooled in research studies. Metrics developed

1 from a large database such as NHANES-II should help with differentiating between fractures and normal  
2 variants of shape. The deterministic thresholds for standardized metrics that can provide clinically  
3 meaningful diagnostic or prognostic information are yet to be determined. These data may also be  
4 useful for quality control in technology designed to automatically obtain landmarks coordinates and  
5 other metrics from medical imaging.

6

## 1 References

- 2 1. Lindsay R, Silverman SL, Cooper C, Hanley DA, Barton I, Broy SB, et al. Risk of new vertebral  
3 fracture in the year following a fracture. *Jama*. 2001;285(3):320-3.
- 4 2. Ross PD, Davis JW, Epstein RS, Wasnich RD. Pre-existing fractures and bone mass predict  
5 vertebral fracture incidence in women. *Ann Intern Med*. 1991;114:919-23.
- 6 3. Lunt M, O'Neill TW, Felsenberg D, Reeve J, Kanis JA, Cooper C, et al. Characteristics of a  
7 prevalent vertebral deformity predict subsequent vertebral fracture: results from the European  
8 Prospective Osteoporosis Study (EPOS). *Bone*. 2003;33(4):505-13.
- 9 4. Gehlbach S, Bigelow C, Heimisdottir M, May S, Walker M, Kirkwood J. Recognition of vertebral  
10 fracture in a clinical setting. *Osteoporosis international*. 2000;11(7):577-82.
- 11 5. Delmas PD, van de Langerijt L, Watts NB, Eastell R, Genant H, Grauer A, et al. Underdiagnosis of  
12 vertebral fractures is a worldwide problem: the IMPACT study. *Journal of bone and mineral  
13 research*. 2005;20(4):557-63.
- 14 6. Ferrar L, Jiang G, Adams J, Eastell R. Identification of vertebral fractures: an update.  
15 *OsteoporosInt*. 2005;16(7):717-28.
- 16 7. Veleanu C, Diaconescu N. Contribution to the clinical anatomy of the vertebral column.  
17 Considerations on the stability and the instability at the height of the "vertebral units".  
18 *Anatomischer Anzeiger*. 1975;137(3):287-95.
- 19 8. Saraste H, Brostrom LA, Aparisi T, Axdorph G. Radiographic measurement of the lumbar spine. A  
20 clinical and experimental study in man. *Spine*. 1985;10(3):236-41.
- 21 9. Frymoyer JW, Phillips RB, Newberg AH, MacPherson BV. A comparative analysis of the  
22 interpretations of lumbar spinal radiographs by chiropractors and medical doctors. *Spine*.  
23 1986;11(10):1020-3.
- 24 10. Gallagher JC, Hedlund LR, Stoner S, Meeger C. Vertebral morphometry: normative data. *Bone  
25 Miner*. 1988;4(2):189-96.
- 26 11. Davies KM, Recker RR, Heaney RP. Normal vertebral dimensions and normal variation in serial  
27 measurements of vertebrae. *J Bone Min Res*. 1989;4:341-9.
- 28 12. Smith-Bindman R, Cummings SR, Steiger P, Genant HK. A comparison of morphometric  
29 definitions of vertebral fracture. *JBone MinerRes*. 1991;6(1):25-34.
- 30 13. Wang TM, Shih C. Morphometric variations of the lumbar vertebrae between Chinese and  
31 Indian adults. *Acta Anat(Basel)*. 1992;144(1):23-9.
- 32 14. Nicholson PH, Haddaway MJ, Davie MW, Evans SF. Vertebral deformity, bone mineral density,  
33 back pain and height loss in unscreened women over 50 years. *Osteoporosis International*. Dec  
34 1993;3(6):300-7.
- 35 15. O'Neill T, Varlow J, Felsenberg D, Johnell O, Weber K, Marchant F, et al. Variation in vertebral  
36 height ratios in population studies. *J Bone Miner Res*. 1994;12:1895-907.
- 37 16. Brinckmann P, Frobin W, Biggemann M, Hilweg D, Seidel S, Burton K, et al. Quantification of  
38 overload injuries to thoracolumbar vertebrae and discs in persons exposed to heavy physical  
39 exertions or vibration at the work-place: The shape of vertebrae and intervertebral discs—study  
40 of a young, healthy population and a middle-aged control group. *Clinical Biomechanics*.  
41 1994;9:S3-S83.
- 42 17. Black DM, Palermo L, Nevitt MC, Genant HK, Epstein R, San VR, et al. Comparison of methods for  
43 defining prevalent vertebral deformities: the Study of Osteoporotic Fractures. *JBone MinerRes*.  
44 1995;10(6):890-902.



- 1 18. Wu CY, Li J, Jergas M, Genant HK. Comparison of semiquantitative and quantitative techniques  
2 for the assessment of prevalent and incident vertebral fractures. *OsteoporosInt*. 1995;5(5):354-  
3 70.
- 4 19. Genant HK, Jergas M, Palermo L, Nevitt M, Valentin RS, Black D, et al. Comparison of  
5 semiquantitative visual and quantitative morphometric assessment of prevalent and incident  
6 vertebral fractures in osteoporosis The Study of Osteoporotic Fractures Research Group. *JBone*  
7 *MinerRes*. 1996;11(7):984-96.
- 8 20. Frobin W, Brinckmann P, Biggemann M, Tillotson M, Burton K. Precision measurement of disc  
9 height, vertebral height and sagittal plane displacement from lateral radiographic views of the  
10 lumbar spine. *Clinical Biomechanics*. 1997;12:S1-S63.
- 11 21. Genant HK. Assessment of vertebral fractures in osteoporosis research. *JRheumatol*.  
12 1997;24(6):1212-4.
- 13 22. Brinckmann P, Frobin W, Biggemann M, Tillotson M, Burton K, Burke C, et al. Quantification of  
14 overload injuries to thoracolumbar vertebrae and discs in persons exposed to heavy physical  
15 exertions or vibration at the workplace Part II Occurrence and magnitude of overload injury in  
16 exposed cohorts. *Clinical Biomechanics*. 1998;13:S1-S36.
- 17 23. Espeland A, Korsbrekke K, Albrektsen G, Larsen JL. Observer variation in plain radiography of the  
18 lumbosacral spine. *BrJ Radiol*. 1998;71(844):366-75.
- 19 24. Ferrar L, Eastell R. Identification of vertebral deformities in men: comparison of morphometric  
20 radiography and morphometric X-ray absorptiometry. *OsteoporosInt*. 1999;10(2):167-74.
- 21 25. Black DM, Palermo L, Nevitt MC, Genant HK, Christensen L, Cummings SR. Defining incident  
22 vertebral deformity: a prospective comparison of several approaches. The Study of Osteoporotic  
23 Fractures Research Group. *JBone MinerRes*. 1999;14(1):90-101.
- 24 26. Weber K, Lunt M, Gowin W, Lauer mann T, Armbrecht G, Wieland E, et al. Measurement  
25 imprecision in vertebral morphometry of spinal radiographs obtained in the European  
26 Prospective Osteoporosis Study: consequences for the identification of prevalent and incident  
27 deformities. *BrJ Radiol*. 1999;72(862):957-66.
- 28 27. Zhou SH, McCarthy ID, McGregor AH, Coombs RR, Hughes SP. Geometrical dimensions of the  
29 lower lumbar vertebrae--analysis of data from digitised CT images. *Eur Spine J*. 2000;9(3):242-8.
- 30 28. Szulc P, Munoz F, Sornay-Rendu E, Paris E, Souhami E, Zanchetta J, et al. Comparison of  
31 morphometric assessment of prevalent vertebral deformities in women using different  
32 reference data. *Bone*. 2000;27(6):841-6.
- 33 29. Ferrar L, Jiang G, Barrington NA, Eastell R. Identification of vertebral deformities in women:  
34 comparison of radiological assessment and quantitative morphometry using morphometric  
35 radiography and morphometric X-ray absorptiometry. *JBone MinerRes*. 2000;15(3):575-85.
- 36 30. Rea JA, Chen MB, Li J, Blake GM, Steiger P, Genant HK, et al. Morphometric X-ray  
37 absorptiometry and morphometric radiography of the spine: a comparison of prevalent  
38 vertebral deformity identification. *JBone MinerRes*. 2000;15(3):564-74.
- 39 31. Rea JA, Chen MB, Li J, Marsh E, Fan B, Blake GM, et al. Vertebral morphometry: a comparison of  
40 long-term precision of morphometric X-ray absorptiometry and morphometric radiography in  
41 normal and osteoporotic subjects. *OsteoporosInt*. 2001;12(2):158-66.
- 42 32. Ferrar L, Jiang G, Eastell R. Vertebral wedge angle measured by morphometric X-ray  
43 absorptiometry. *OsteoporosInt*. 2001;12(11):914-21.
- 44 33. Szulc P, Munoz F, Marchand F, Delmas PD. Semiquantitative evaluation of prevalent vertebral  
45 deformities in men and their relationship with osteoporosis: the MINOS study. *OsteoporosInt*.  
46 2001;12(4):302-10.

- 1 34. Frobin W, Leivseth G, Biggemann M, Brinckmann P. Vertebral height, disc height,  
2 posteroanterior displacement and dens-atlas gap in the cervical spine: precision measurement  
3 protocol and normal data. *ClinBiomech(Bristol, Avon)*. 2002;17(6):423-31.
- 4 35. Shao Z, Rompe G, Schiltenswolf M. Radiographic changes in the lumbar intervertebral discs and  
5 lumbar vertebrae with age. *Spine*. 2002;27(3):263-8.
- 6 36. Pfeiffer M, Geisel T. Analysis of a computer-assisted technique for measuring the lumbar spine  
7 on radiographs: correlation of two methods. *AcadRadiol*. 2003;10(3):275-82.
- 8 37. Keynan O, Fisher CG, Vaccaro A, Fehlings MG, Oner FC, Dietz J, et al. Radiographic measurement  
9 parameters in thoracolumbar fractures: a systematic review and consensus statement of the  
10 spine trauma study group. *Spine (Phila Pa 1976)*. 10.1097/01.brs.0000201261.94907.0d doi  
11 ;00007632-200603010-00025 pii 2006;31(5):E156-E65.
- 12 38. Fink HA, Milavetz DL, Palermo L, Nevitt MC, Cauley JA, Genant HK, et al. What proportion of  
13 incident radiographic vertebral deformities is clinically diagnosed and vice versa? *JBone  
14 MinerRes*. 2005;20(7):1216-22.
- 15 39. O'Neill TW, Varlow J, Felsenberg D, Johnell O, Weber K, Marchant F, et al. Variation in vertebral  
16 height ratios in population studies. European Vertebral Osteoporosis Study Group. *Journal of  
17 Bone & Mineral Research*. Dec 1994;9(12):1895-907.
- 18 40. Schwartz EN, Steinberg D. Detection of vertebral fractures. *CurrOsteoporosRep*. 2005;3(4):126-  
19 35.
- 20 41. Pflugmacher R, Schroeder RJ, Klostermann CK. Incidence of adjacent vertebral fractures in  
21 patients treated with balloon kyphoplasty: two years' prospective follow-up. *Acta Radiol*.  
22 P22233334L147355 pii ;10.1080/02841850600854928 doi 2006;47(8):830-40.
- 23 42. Shindle MK, Gardner MJ, Koob J, Bukata S, Cabin JA, Lane JM. Vertebral height restoration in  
24 osteoporotic compression fractures: kyphoplasty balloon tamp is superior to postural correction  
25 alone. *Osteoporosis International*. 2006;17(12):1815-9.
- 26 43. Siminoski K, Jiang G, Adachi J, Hanley D, Cline G, Ioannidis G, et al. Accuracy of height loss during  
27 prospective monitoring for detection of incident vertebral fractures. *Osteoporosis International*.  
28 2005;16(4):403-10.
- 29 44. Zhao KD, Ben-Abraham EI, Magnuson DJ, Camp JJ, Berglund LJ, An KN, et al. Effect of Off-Axis  
30 Fluoroscopy Imaging on Two-Dimensional Kinematics in the Lumbar Spine: A Dynamic In Vitro  
31 Validation Study. *J Biomech Eng*. May 2016;138(5):054502.
- 32 45. Zhao KD, Yang C, Zhao C, Stans AA, An KN. Assessment of noninvasive intervertebral motion  
33 measurements in the lumbar spine. *JBiomechanics*. 2005;38(9):1943-6.
- 34 46. Reitman CA, Hipp JA, Nguyen L, Esses SI. Changes in segmental intervertebral motion adjacent to  
35 cervical arthrodesis: a prospective study. *Spine (Phila Pa 1976)*. Jun 1 2004;29(11):E221-6. Epub  
36 2004/05/29.
- 37 47. Reitman CA, Mauro KM, Nguyen L, Ziegler JM, Hipp JA. Intervertebral motion between flexion  
38 and extension in asymptomatic individuals. *Spine (Phila Pa 1976)*. Dec 15 2004;29(24):2832-43.  
39 Epub 2004/12/16.
- 40 48. Patwardhan AG, Havey RM, Wharton ND, Tsitsopoulos PP, Newman P, Carandang G, et al.  
41 Asymmetric motion distribution between components of a mobile-core lumbar disc prosthesis:  
42 an explanation of unequal wear distribution in explanted CHARITE polyethylene cores. *J Bone  
43 Joint Surg Am*. May 2 2012;94(9):846-54.
- 44 49. Pearson AM, Spratt KF, Genuario J, McGough W, Kosman K, Lurie J, et al. Precision of lumbar  
45 intervertebral measurements: Does a computer-assisted technique improve reliability? *Spine*.  
46 2011;36(7):572-80.
- 47 50. Hurxthal LM. Measurement of anterior vertebral compressions and biconcave vertebrae. *Am J  
48 Roentgenol*. 1968;103:635-44.

- 1 51. Diacinti D, Guglielmi G. Vertebral morphometry. *Radiologic Clinics*. 2010;48(3):561-75.
- 2 52. Genant HK, Wu CY, van KC, Nevitt MC. Vertebral fracture assessment using a semiquantitative  
3 technique. *JBone MinerRes*. 1993;8(9):1137-48.
- 4 53. CDC. Examination Staff Procedures Manual for the Health and Nutrition Examination Survey,  
5 1976-1979. Centers for Disease Control and Prevention; 1976.
- 6 54. Black DM, Cummings SR, Stone K, Hudes E, Palermo L, Steiger P. A new approach to defining  
7 normal vertebral dimensions. *J Bone Min Res*. 1991;6:883-92.
- 8 55. Diacinti D, Pisani D, Del Fiacco R, Francucci CM, Fiore CE, Frediani B, et al. Vertebral  
9 morphometry by X-ray absorptiometry: which reference data for vertebral heights? *Bone*.  
10 2011;49(3):526-36.
- 11 56. Djoumessi RMZ, Maalouf G, Wehbe J, Nehme A, Maalouf N, Seeman E. The varying distribution  
12 of intra-and inter-vertebral height ratios determines the prevalence of vertebral fractures. *Bone*.  
13 2004;35(2):348-56.
- 14 57. Rea J, Steiger P, Blake G, Potts E, Smith I, Fogelman I. Morphometric X-ray absorptiometry:  
15 reference data for vertebral dimensions. *Journal of Bone and Mineral Research*. 1998;13(3):464-  
16 74.
- 17 58. Staub BN, Holman PJ, Reitman CA, Hipp JA. Sagittal plane lumbar intervertebral motion during  
18 seated flexion-extension in 658 asymptomatic, non-degenerated levels. *J Neurosurgery Spine*.  
19 2015 2015;23(6):731-8.
- 20 59. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an  
21 image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging*.  
22 2012;30(9):1323-41.
- 23 60. Shackelford JA, Shusharina N, Verberg J, Warmerdam G, Winey B, Neuner M, et al. Plastimatch  
24 1.6-current capabilities and future directions. *MICCAI, First International Workshop on Image-  
25 Guidance and Multimodal Dose Planning in Radiation Therapy*2012.
- 26 61. Antani S, Cheng J, Long J, Long LR, Thoma GR. Medical validation and CBIR of spine x-ray images  
27 over the Internet. *Internet Imaging VII: International Society for Optics and Photonics*; 2006. p.  
28 60610J.
- 29 62. Goh S, Tan C, Price R, Edmondston S, Song S, Davis S, et al. Influence of age and gender on  
30 thoracic vertebral body shape and disc degeneration: an MR investigation of 169 cases. *The  
31 Journal of Anatomy*. 2000;197(4):647-57.
- 32 63. Cheng X, Sun Y, Boonen S, Nicholson P, Brys P, Dequeker J, et al. Measurements of vertebral  
33 shape by radiographic morphometry: sex differences and relationships with vertebral level and  
34 lumbar lordosis. *Skeletal radiology*. 1998;27(7):380-4.
- 35 64. McCloskey E, Spector T, Eyres K, Fern E, O'Rourke N, Vasikaran S, et al. The assessment of  
36 vertebral deformity: a method for use in population studies and clinical trials. *Osteoporosis Int*.  
37 1993;3:138-47.
- 38 65. Fukunaga M, Nakamura T, Shiraki M, Kuroda T, Ohta H, Hosoi T, et al. Absolute height reduction  
39 and percent height ratio of the vertebral body in incident fracture in Japanese women. *Journal  
40 of bone and mineral metabolism*. 2004;22(2):104-10.
- 41 66. Lau E, Chan H, Woo J, Lin F, Black D, Nevitt M, et al. Normal ranges for vertebral height ratios  
42 and prevalence of vertebral fracture in Hong Kong Chinese: a comparison with American  
43 Caucasians. *Journal of Bone and Mineral Research*. 1996;11(9):1364-8.
- 44 67. Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ, III. Classification of vertebral fractures. *J  
45 Bone Min Res*. 1991;6:207-15.
- 46 68. O'Brien MF, Kuklo TR, Blanke KM, Lenke LG. Spinal deformity study group radiographic  
47 measurement manual. Memphis, TN: Medtronic Sofamor Danek. 2005.

- 1 69. Lian J, Levine N, Cho W. A review of lumbosacral transitional vertebrae and associated vertebral  
2 numeration. *European Spine Journal*. 2018;27(5):995-1004.
- 3 70. Shah M, Halalmeh DR, Sandio A, Tubbs RS, Moisi MD. Anatomical variations that can lead to  
4 spine surgery at the wrong level: Part III lumbosacral spine. *Cureus*. 2020;12(7).
- 5 71. Abola MV, Teplensky JR, Cooperman DR, Bauer JM, Liu RW. Pelvic Incidence in Spines With 4  
6 and 6 Lumbar Vertebrae. *Global Spine Journal*. 2019;9(7):708-12.
- 7 72. Hanhivaara J, Määttä JH, Niinimäki J, Nevalainen MT. Lumbosacral transitional vertebrae are  
8 associated with lumbar degeneration: retrospective evaluation of 3855 consecutive abdominal  
9 CT scans. *European Radiology*. 2020;30(6):3409-16.
- 10 73. Apazidis A, Ricart PA, Diefenbach CM, Spivak JM. The prevalence of transitional vertebrae in the  
11 lumbar spine. *The Spine Journal*. 2011;11(9):858-62.
- 12 74. Bron JL, van Royen BJ, Wuisman P. The clinical significance of lumbosacral transitional  
13 anomalies. *Acta Orthopaedica Belgica*. 2007;73(6):687.
- 14 75. Konin G, Walz D. Lumbosacral transitional vertebrae: classification, imaging findings, and clinical  
15 relevance. *American Journal of Neuroradiology*. 2010;31(10):1778-86.
- 16 76. Tang M, Yang X-f, Yang S-w, Han P, Ma Y-m, Yu H, et al. Lumbosacral transitional vertebra in a  
17 population-based study of 5860 individuals: Prevalence and relationship to low back pain.  
18 *European journal of radiology*. 2014;83(9):1679-82.
- 19 77. Nardo L, Alizai H, Virayavanich W, Liu F, Hernandez A, Lynch JA, et al. Lumbosacral transitional  
20 vertebrae: association with low back pain. *Radiology*. 2012;265(2):497-503.
- 21 78. Ravi B, Rampersaud R. Clinical magnification error in lateral spinal digital radiographs. *Spine*.  
22 2008;33(10):E311-E6.
- 23 79. Shigematsu H, Koizumi M, Yoneda M, Iida J, Oshima T, Tanaka Y. Magnification error in digital  
24 radiographs of the cervical spine against magnetic resonance imaging measurements. *Asian  
25 spine journal*. 2013;7(4):267.
- 26 80. Lunt M, Gowin W, Johnell, Armbrecht G, Felsenberg D. A statistical method to minimize  
27 magnification errors in serial vertebral radiographs. *OsteoporosInt*. 2001;12(11):909-13.

28