

Factors in measuring lumbar spondylolisthesis with reference data from NHANES-II

**John A. Hipp, PhD^{1*}, Trevor F. Grieco¹, PhD, Patrick Newman, MS¹,
Vikas V. Patel², MD, Charles A. Reitman, MD³**

¹ Medical Metrics, Imaging Core Laboratory, Houston, TX 77056

² Colorado University Orthopedics

³ Orthopaedics and Physical Medicine, Medical University of South Carolina, Charleston, SC 29425

Keywords: Spondylolisthesis, disc height, disc angle, normal, radiographic

* Corresponding Author: John A. Hipp, PhD

Email address: jhipp@medicalmetrics.com

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

Phone: (360) 402-0022

Acknowledgements: The publicly available NHANES-II database with images is gratefully acknowledged. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

https://www.cdc.gov/nchs/data/nhanes/nhanes_release_policy.pdf

Dislosures: JAH, TFG, and PN are employees of Medical Metrics, Inc which is an imaging core laboratory with the technology used to make all of the radiographic measurements. They each received salary support to complete the work. No external funding was received in support of this work.

Abstract

Background

Although spondylolisthesis, disc height loss, and abnormal disc angles are commonly assessed for in clinical practice and research studies, the factors that influence measurements are incompletely understood, and reference data representing a very large and diverse population are not available. Over 7000 lumbar spine x-rays were collected as part of the 2nd National Health and Nutrition Examination Survey (NHANES-II). This nationwide probability sample may facilitate development of robust reference data to objectively classify spondylolisthesis, disc heights, or disc angles as normal vs abnormal. Analysis of lumbar spine x-rays may also help to document whether disc heights and disc angles must be considered when assessing for spondylolisthesis.

Methods

Anatomic landmarks were obtained using neural networks and coded logic for L1 to S1 in the NHANES-II lumbar spine radiographs. Nine disc and sagittal plane offset (SPO) measurements were calculated from the landmarks. These data were trimmed to exclude abnormal discs and SPO. The factors that affect SPO were explored along with associations between the metrics and with age, sex, race, nation of origin and BMI. The prevalence of abnormalities was tabulated. Metrics were also calculated for previously analyzed lumbar flexion-extension x-rays to better understand the dependence of SPO on disc angles and disc heights. The errors that occur due to variability in radiographic projection were also assessed.

Results

SPO and disc metrics were obtained for 35,490 levels from 7,358 NHANES-II subjects who were 25 to 74 years old. Descriptive statistics for nine SPO and disc metrics were tabulated before and after trimming the data. Age, BMI, and sex were statistically significant but explained little of the variance in the metrics. SPO was significantly dependent on disc angle and height, though less so in the NHANES radiographs than in lumbar flexion-extension studies. Errors in the metrics due to out-of-plane imaging are generally small but can be large with extensive out-of-plane x-rays.

Discussion

The NHANES-II collection of lumbar x-rays allows for establishing reference data for SPO and disc metrics. These reference data allow for easily interpreted standardized reporting in units of std dev from average normal. SPO was significantly dependent on disc angle and disc height although the effect is small when there is limited flexion or extension. If SPO is being assessed from flexion or extension, a simple correction can be made. Caution is needed when interpreting measurements when the x-ray beam passes obliquely through the endplates or posterior wall (out-of-plane imaging).

Conclusions

The NHANES-II data may facilitate standardized assessments of SPO, disc heights and disc angles. Adjustments should be applied to SPO measurements if made from x-rays with the patient in flexed or extended positions.

1 Factors in measuring lumbar 2 spondylolisthesis with reference data from 3 NHANES-II 4

5 Introduction

6 Spondylolisthesis is defined as an abnormal “slip” of one vertebra relative to another and is most
7 commonly assessed in the sagittal plane.¹ Some amount of “slip” occurs with healthy discs. The challenge
8 is to reliably classify “slip” as normal vs abnormal while accounting for significant co-factors. That requires
9 both a reliable measurement method, reference data defining normal vs abnormal, and an understanding of
10 important cofactors. In clinical practice, spondylolisthesis is often subjectively assessed, due to the time
11 required to carefully make the measurements, and the lack of definitive evidence that quantitative
12 measurements are needed. Definitive evidence for/against quantitative measurements requires yet-to-be-
13 completed research, and the outcome will depend on the quality of the quantitative measurements and
14 reference data. Adoption and utilization in routine clinical practice will also depend on the efficiency with
15 which such measurements can be made.

16 With the goal of reserving the use of the term “spondylolisthesis” for abnormal sagittal plane
17 positioning between vertebrae, the term “sagittal plane offset” (SPO) will be used generically for all
18 measurements of the offset of one vertebra with respect to the adjacent vertebra. It is assumed that a range
19 of SPO will be “normal” and SPO outside of that range can be labeled spondylolisthesis (either anterior or
20 posterior). When SPO or disc height loss is pronounced, the reliability of current clinical assessment may
21 be sufficient with current methods, but when it is more subtle, reliable reference data are needed as well
22 as methods to account for confounding variables. It is also possible that new insights can be gained from

¹ <http://www.wheelsonline.com/ISSLS/section-15-chapter-2-radiology-and-degenerative-spondylolisthesis/>

1 using a continuous quantitative SPO metric in research studies. The challenge is to validate criteria for
2 classifying “normal” versus “abnormal” SPO, that account for important variables and sources of error.
3 Objective criteria for classifying disc heights or disc angles as normal vs abnormal are also poorly
4 developed and additional reference data are needed, as evidenced by the lack of standardization in how
5 SPO is measured and reported in research and clinical practice.

6 The rapid advancement of artificial intelligence/machine learning technology in the field of spine
7 care will make near instant quantitative measurements of spondylolisthesis, disc heights, angles and other
8 metrics possible¹⁻³. That technology will require validated guidelines for how to optimally interpret these
9 measurements. Once guidelines are validated, evidence can be generated to determine if and how automated
10 quantitative measurements might improve outcomes and lower costs for patients with symptoms or
11 treatment outcomes that may be related to abnormal SPO, disc heights or disc angles. The value of a
12 quantitative diagnostic is in part determined by measurement errors, so understanding how to minimize
13 sources of error in spondylolisthesis and disc measurements is a necessary prerequisite to effective clinical
14 trials.

15 Multiple factors may influence SPO measurements. For example, SPO is dependent on patient
16 positioning⁴. This may be due to the differences in spinal loading conditions between positions but may
17 also be partially due to how SPO measurements are affected by disc angle and disc height. An objective
18 metric that can account for variability in disc angle and disc height, while also accounting for differences
19 between levels and sexes and other variables, may facilitate more precise diagnosis as well as allow for
20 more sensitive detection of changes over time. If accurate measurement of SPO must account for multiple
21 factors, it is possible this can be automated and accomplished algorithmically with reference data from a
22 reliable analysis of a large population.

23

1 Hypotheses

- 2 1. SPO is dependent on level (eg L1-L2 vs L4-L5), disc height and disc angle, as well as other
3 variables.
- 4 2. If hypothesis 1 is true, then an objective metric to quantify SPO can be developed that accounts
5 for: 1) the variability in disc angle due to patient positioning, 2) the normal variability in disc
6 height between patients, 3) the variability in disc height between levels, and 4) the variability of
7 normal SPO as it exists in asymptomatic volunteers at normal levels.
- 8 3. The above two hypotheses can be tested using the database of over 7,000 lumbar spine
9 radiographs from the NHANES-II study, supplemented with a collection of flexion-extension
10 radiographs to better understand the importance of disc angle.
- 11 4. SPO and disc metrics will depend in part on variability in radiographic projection and this
12 phenomena can be understood by analysis of precisely calculated landmarks obtained from
13 variable, digitally reconstructed radiographic projections.

14

15 NHANES-II

16

17 Between 1976 and 1980, the 2nd National Health and Nutrition Examination Survey (NHANES-
18 II) was conducted.⁵ This was a nationwide probability sample to document the health status of the United
19 States. Justified by the prevalence and societal impact of back pain, approximately 7,000 lateral lumbar
20 spine X-rays were collected as part of this survey. One lateral lumbar X-ray was obtained for each
21 participant. Demographic, anthropometric, health, and medical history data were also collected.
22 NHANES-II may thereby be useful to establish normative reference data for use in clinical practice and
23 research studies.

24

1 Materials and Methods

2 The NHANES-II images and data were obtained through public access². A previously validated
3 series of neural networks and coded logic (Spine CAMP™, Medical Metrics, Inc., Houston, TX) were
4 used to automatically obtain four landmarks for each vertebral body from L1 to S1 (Supplemental
5 Appendix 1, Figure 1). Although the NHANES-II lumbar X-rays almost always included up to T10 in the
6 field-of-view, the neural networks were only trained to recognize L1 to S1. The series of neural networks
7 include quality control networks to assure the X-ray was a lateral lumbar view, manipulate the image if
8 needed so upper vertebrae are toward the top of the X-ray, assure spinous processes toward the left side of
9 the image, and assure that bone appears whiter than air. Neural networks and coded logic were then used
10 to segment the bone, find individual vertebrae, label the vertebrae, find the four corners of each vertebra,
11 and finally refrain from reporting landmark coordinates when the confidence of the networks was too low.
12 In addition, an anomaly detection algorithm was used to identify likely errors in landmark placement.
13 Those levels were excluded from the data analysis. All the neural networks were trained with over twenty
14 thousand lateral X-rays where analysts had previously digitized standardized landmarks. The NHANES-II
15 images were not used in training the neural networks. The analysts had used Quantitative Motion
16 Analysis software (QMA®, Medical Metrics, Inc, Houston, TX) to place the landmarks. The QMA®
17 software has previously been validated.⁶⁻¹¹ QMA® has been used to produce data reported in over 170
18 peer-reviewed spine studies (references available upon request). Vertebral labeling in the data used to
19 train the neural networks was based on identifying the vertebra most characteristic of S1 and then labeling
20 other vertebrae relative to S1.

21 Details of landmark placement are provided in Supplemental Appendix 1, since all SPO and disc
22 measurements will be dependent on exactly how the landmarks are placed¹². Vertebral morphology
23 calculated from the landmarks has been previously reported.¹³

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

1 The following nine metrics were calculated from the landmarks for each intervertebral level from L1-
2 L2 to L5-S1, using one of two methods:

3 • Method 1

4 ○ Anterior disc height (ADH)

5 ○ Posterior disc height (PDH)

6 ○ Disc angle (DA)

7 ○ Anterior SPO (ASPO)

8 ○ Posterior SPO (PSPO)

9 • Method 2

10 ○ Ventral disc height (VDH)

11 ○ Dorsal disc height (DDH)

12 ○ Mid-plane angle (MPA)

13 ○ Centroid SPO (CSPO)

14 The average disc height was also calculated from ADH and PDH as that may be of value in some
15 applications. The disc area was also calculated (in units of endplate width squared) from the four
16 landmarks using the shoelace algorithm, as this may be of interest in some applications.

17 Method 1 metrics are illustrated in Figure . Method 2 metrics were all first described by Frobin et
18 al and are illustrated in Figures 2 and 3.^{14,15} Method 2 was developed by Frobin et al with the goal of
19 minimizing the influence of radiographic distortion on measurements. The ventral and dorsal terminology
20 used by Frobin et al is retained in this paper to help distinguish between height metrics calculated using
21 method 1 versus method 2. Since substantial variability can occur in radiographic magnification¹⁶⁻¹⁸, and
22 since there was no scaling device in the NHANES-II X-rays, all SPO and disc height metrics were
23 calculated in units of % endplate width, using the superior endplate of the inferior vertebra. Note that
24 interpretation of the sign of spondylolisthesis is based on local anatomy and not global anatomy. A
25 positive ASPO informs that the anterior inferior corner of the superior vertebra is anterior to the anterior

1 superior corner of the inferior vertebra. A positive PSPO informs that the posterior inferior corner of the
2 superior vertebra is posterior to the posterior superior corner of the inferior vertebra. A negative DA or
3 MPA informs that the ADH (or VDH) is smaller than the PDH (or DDH).

4 With the goal of establishing normal SPO, disc heights, and disc angles, abnormalities in the
5 NHANES-II database were trimmed based on the distribution of the data. The goal was to achieve
6 normally distributed data that describe healthy disc metrics and SPO, so outliers were trimmed based on
7 percentiles in the distributions. Multivariate analysis of variance was used to understand associations
8 between the SPO and disc metrics, and age, sex, BMI and other variables. All statistical analysis was
9 completed using Stata ver 15 (College Station, TX).

10 Since all the NHANES-II radiographs were obtained with the subject side-lying, there was
11 generally not a lot of variation in the flexion angle relative to what can occur in clinical practice. To better
12 understand the effect of disc angle and disc heights on SPO, flexion-extension radiographs from a prior
13 study¹⁹ were also analyzed to obtain a larger range of disc angles. Regression analysis with post-
14 estimation calculations was used to establish coefficients for equations that allow for predicting normal
15 SPO from the flexion-extension radiographs based on the most predictive variables, excluding the
16 outliers. This is described in Supplemental Appendix 2.

17 It is known that variability in radiographic projection can affect measurements of the spine.^{15,20}
18 To help appreciate the magnitude of error that can occur in the nine SPO and disc metrics, simulated X-
19 rays were used that were generated with precise knowledge of the true location of the landmarks on the
20 X-ray. This experiment is described in Supplemental Appendix 3.

21

22

23

1 Results

3 Summary of Data Analyzed

4 SPO and disc metrics were obtained for 35,446 levels from 7,358 NHANES-II subjects who were
5 25 to 74 years old. It took approximately 9 hrs on a dual CPU, 4 GPU (Nvidia T4) production server to
6 generate the landmarks for all X-rays, that were then used to calculate the disc and SPO metrics. Data
7 were not produced for 1300 levels (3.5% of total possible levels) due to issues detected by the neural
8 networks and coded logic. An additional 44 levels (0.12%) were excluded by anomaly detection. The age
9 and body size data are summarized in Table 1. Older ages are disproportionately represented (Figure 4).

10

11 **Table 1:** Sample size, average age and [SD], and average BMI (kg/m²) and [SD] for the
12 NHANES-II subjects that were analyzed.

Sex	N	Age	BMI
Male	4,552	50.9[15.3]	25.5[4.0]
Female	2,806	63.3[6.4]	26.4[5.5]

13

14

15 Correlations between disc and SPO metrics

16

17 The method 1 and method 2 disc and SPO metrics were significantly correlated with each other
18 (Pearson correlation coefficients between each pair of variables, for the L3L4 level, are provided in Table
19 2). Correlations were nearly perfect between anterior disc height and ventral disc height, between
20 posterior disc height and dorsal disc height, and between average disc height and disc area. Disc angle
21 (method 1) and mid-plane disc angle (method 2) were also strongly correlated. Note that ASPO and

1 CSPO are positively correlated while PSPO and CSPO are negatively correlated since PSPO quantifies
 2 the amount of translation of the posterior-inferior corner of the superior vertebra into the spinal canal.
 3
 4 Table 2: Pearson correlation coefficients between each pair of variables, for the L3L4 level, including
 5 data for all subjects. Every correlation coefficient is statistically significant in this table ($P < 0.0001$).
 6 Similar results were found at all other levels.

	ADH	PDH	Avg DH	DA	ASPO	PSPO	VDH	DDH	MPA	CSPO	Disc AREA
ADH	1										
PDH	0.46	1									
AvgDH	0.90	0.80	1								
DA	0.70	-0.30	0.32	1							
ASPO	-0.31	0.20	-0.11	-0.52	1						
PSPO	0.29	-0.07	0.16	0.34	-0.74	1					
VDH	0.99	0.43	0.87	0.73	-0.39	0.27	1				
DDH	0.48	0.99	0.81	-0.27	0.13	-0.05	0.46	1			
MPA	0.53	-0.32	0.20	0.82	-0.36	0.19	0.55	-0.30	1		
CSPO	-0.15	0.11	-0.05	-0.26	0.80	-0.83	-0.18	0.07	-0.17	1	
Disc AREA	0.91	0.78	1.00	0.35	-0.13	0.23	0.88	0.79	0.21	-0.09	1

7

8 [Establishing normal disc and SPO metrics](#)

9

10 With the goal of using the NHANES-II X-rays to develop reference data that can define “normal”
 11 disc heights, angles, and SPO, outliers were trimmed based on percentiles. Initial assessment of the
 12 distributions of disc metrics, including all NHANES-II subjects, revealed that every metric was
 13 significantly skewed with significant kurtosis ($p < 0.0001$ for both skewness and kurtosis, for every metric
 14 at almost every level). One set of trimming criteria (such as excluding the bottom and top 25%) was not

1 used for all metrics, because the distribution of data post-trimming was clearly not normal for all metrics.
2 This is understandable, since for example, disc height loss from degeneration is far more likely than
3 abnormal disc space widening. If too much of the tails of the entire dataset were removed for any metric,
4 the tails of the distribution were clearly truncated (not a gaussian distribution) and/or the distribution was
5 skewed. Table 3 describes the criteria used to trim outliers for each of the variables. The proportions of
6 the upper and lower tails that were trimmed per Table 3 were selected to achieve skewness close to 0 and
7 kurtosis close to 3 (ideal for a normal distribution), although a formal optimization scheme was not used.

8 Table 3: criteria for identifying and trimming outliers in the data.

Metric	Acronym	Lower Proportion Removed	Upper Proportion Removed
Method 1			
Anterior Disc Height	ADH	10	5
Posterior Disc Height	PDH	10	5
Disc Angle	DA	10	10
Anterior Sagittal Plane Offset	ASPO	5	5
Posterior Sagittal Plane Offset	PSPO	5	5
Method 2			
Ventral Disc Height	VDH	10	5
Dorsal Disc Height	DDH	10	5
MidPlane Disc Angle	MPA	10	10
Centroid Sagittal Plane Offset	CSPO	5	5

9
10 Supplemental Appendix 1, Tables 1 to 10, provide descriptive statistics for the NHANES-II
11 lumbar SPO and disc metrics, before and after trimming the data. Note that trimming: 1) eliminates
12 approximately half of the data, since if any of the nine metrics was an outlier (based on Table 3), all data

1 for that level was excluded; 2) does not substantially change the mean or median; 3) substantially lowers
2 the coefficient of variation; 4) brings skewness close to 0 and kurtosis closer to 3.

3 4 [Analysis of sources of variance in SPO metrics](#) 5

6 Based on multivariate analysis of variance (ANOVA), after excluding trimmed data (per criteria
7 in Table 3), ASPO and PSPO (method 1) were significantly dependent on level, anterior and posterior
8 disc heights, disc angle, age, and BMI ($P < 0.0001$ for all variables). Sex, race, and nation of origin were
9 not significant for ASPO ($P > 0.78$ for all three). Based on the F-value, disc angle had the strongest
10 association with ASPO, followed by anterior and posterior disc height. The F-value for level and age was
11 an order of magnitude less than for disc angle and disc heights. Sex had a significant association with
12 PSPO, though the F-value was very small. Based on F-values, disc heights and disc angle had the
13 strongest association with PSPO with the F-value for age and level an order of magnitude lower.
14 However, the ANOVA R^2 was 0.56 for ASPO and 0.44 for PSPO, and that would likely not be sufficient
15 to facilitate reliably predicting normal SPO from disc angle and disc height. Too much of the variability
16 would be left unexplained.

17 Based on multivariate analysis of variance (ANOVA), after excluding abnormal measurements
18 (per criteria in Table 3), CSPO (method 2) was significantly dependent on level, ventral and dorsal disc
19 heights, age, BMI ($P < 0.0001$), sex ($P = 0.0002$), and on midplane angle ($P = 0.006$). Race and nation of
20 origin were not significant ($P > 0.83$). Intervertebral level was over an order of magnitude more important
21 than any other variable. The F-value for midplane angle, BMI and sex were small. Including only level,
22 $R^2 = 0.76$ and this was only slightly lower than when all variables were included ($R^2 = 0.77$). Thus, with
23 CSPO, only level is needed to predict normal CSPO, and almost 80% of the variability in CSPO can be
24 explained by level.

25

1 [Applicability of NHANES-II SPO reference data to flexion-extension X-rays](#)

2

3 Supplemental Appendix 2 describes the results of applying the NHANES-II SPO reference data
4 to flexion-extension studies of asymptomatic volunteers. Applying the normative data from the
5 NHANES-II study (where all subjects were imaged in a side-lying position) to 373 flexion-extension
6 radiographs (where subjects were upright), 45% of the flexion and extension radiographs have
7 standardized ASPO metrics >2 or < -2 Std Dev from average normal ASPO. Since 45% of asymptomatic
8 levels can't have abnormal SPO, it is clear that application NHANES-II normative SPO data to flexion
9 and extension X-rays is NOT appropriate.

10 Supplemental appendix 2 also describes development of an equation that can quantify SPO,
11 accounting for disc heights and disc angles, in the presence of flexion and extension. That equation
12 provides the spondylolisthesis index (SI) which is an estimate of the deviation from normal ASPO that
13 accounts for disc heights and disc angle, and can be applied to neutral X-rays as well as flexion-extension
14 X-rays. Applying that equation to the NHANES-II data, 16% of levels had a SI > 2 and 7% had a SI > 3 .
15 Most of the abnormalities were at L4-L5 and L5-S1. The SI and the normalized ASPO metric were
16 linearly correlated ($R^2=0.64$)

17

18 [Prevalence of abnormalities in the NHANES-II lumbar spine radiographs](#)

19

20 Using the post-trimming data in tables 1 to 10 of Supplemental Appendix 1, each measurement in
21 the NHANES-II study was classified as “normal” or “abnormal” (eg ventral disc heights in the lowest 10th
22 percentile were classified as abnormal). The prevalence of abnormalities in the NHANES-II data is
23 provided in Supplementary Appendix 1, Tables 11 to 16. Prevalence was calculated using the
24 standardized (“z”) version of the metrics to simplify interpretation of the data. Both the prevalence of Z
25 scores > 2 and > 3 are provided.

1 Associations with Back Pain

2

3 The subjects in the NHANES II study were asked the question: “Have you ever had pain in your
4 back on most days for at least two weeks”. It is challenging to test for associations between the response
5 to the question and the disc and SPO metrics since it is not clear what level(s) in the spine might be
6 associated with symptoms. Nevertheless, the maximum absolute standardized ASPO, PSPO, and CSPO
7 were calculated for each spine. Logistic regression documents a significant association between
8 maximum absolute standardized deviation from normal SPO and whether the subject reported ever having
9 back pain, $P < 0.05$. However, the odds ratios were close to 1 (1.04 for ASPO and PSPO, and 1.05 for
10 CSPO). Thus, although an association between SPO and back pain can be detected using the NHANES-II
11 data, it is weak. Similarly, the minimum standardized anterior and posterior disc heights were calculated
12 for each subject. Minimum ventral and dorsal disc heights were not calculated since these are nearly
13 perfectly correlated with anterior and posterior disc heights. Both minimum standardized anterior and
14 minimum standardized posterior disc heights were significantly associated with whether the subject
15 reported ever having back pain ($P < 0.0001$). The odds ratios (0.87 for standardized ADH, and 0.85 for
16 standardized PDH) suggest a stronger association between back pain and loss in disc height than found
17 for SPO. The odds ratios from logistic regressions between maximum absolute standardized disc angle
18 and maximum absolute standardized mid-plane angle were also significant, documenting that deviation at
19 any level within a spine from normal disc angle ($P = 0.001$, $OR = 1.09$) or from normal midplane angle
20 ($P = 0.015$, $OR = 1.07$) can be associated with back pain.

21

22

1 **Discussion**

2 A standardized definition of normal SPO and disc metrics, that can be used across research
3 studies and in clinical practice may facilitate advancements in spine care. Analysis of the NHANES data
4 were used to define normal SPO and other disc metrics in an attempt to achieve the goal of a standardized
5 definition. The landmarks used to calculate the metrics were produced using a proprietary
6 network of neural networks and coded logic. However, the landmarks are intended to be
7 equivalent to landmarks produced for many prior research studies. Artificial intelligence has
8 been used to produce vertebral landmarks in many prior studies.⁽⁶²⁻⁶⁷⁾ It is expected that the
9 landmark placement obtained for the NHANES-II radiographs can be reproduced by other
10 methods and the NHANES-II reference data in Supplemental Appendix 1 will thereby be useful
11 in other studies where landmarks are obtained with alternative methods.

12 Further research is needed to determine how to effectively use SPO and other metrics in research
13 and clinical practice. The NHANES-II SPO and disc metrics also document that while some strong
14 correlations exist between a few of these metrics, other metrics are not as strongly correlated supporting
15 that most of the metrics are, in part, detecting unique aspects of disc height, angle, and SPO. The
16 NHANES-II data also support hypothesis 1 in showing that quantitative measurement of SPO are
17 significantly correlated with disc heights and disc angles, though a lot of variability in SPO is not
18 explained by disc heights and disc angles. A prediction equation developed and described in
19 Supplemental Appendix 2 provides an improved method to account for disc heights and disc angles when
20 predicting what should be “normal” SPO for a level. That equation explains more of the variability in
21 SPO.

22 Multiple SPO and disc metrics were investigated. In Figure 1, note that the method 1 anterior and
23 posterior spondylolisthesis measurements are local measurements that quantify the position of a specific
24 corner of the superior vertebra relative to a specific corner of the inferior vertebra, whereas the method 2

1 displacement of the centroid (Figure 3) is a more generic measurement of the position of the superior
2 vertebra relative to the disc bisectrix between superior and inferior vertebrae. One goal of method 2 was
3 to reduce the effects of radiographic distortion, as explained by Frobin et al.¹⁴ The NHANES-II data
4 support that CSPO may be the best metric to use for assessing overall SPO. ASPO and PSPO may prove
5 more effective if the local effects of SPO are important. Carefully contemplating Figures 1 to 3, with
6 respect to how loss of disc height or a change in disc angle might change the SPO measurements, helps to
7 appreciate the hypothesis that disc height and disc angle could change SPO measurements. The research
8 and clinical efficacy of the method 1 versus method 2 SPO measurements remains to be determined.

9 A primary goal was to use the NHANES-II X-rays to establish normal sagittal plane SPO such
10 that abnormal SPO (spondylolisthesis) can be identified and quantified. There currently is no well-
11 validated “gold standard” metric for abnormal versus normal SPO. Insufficient resources prevented
12 manual clinician assessment of all NHANES-II X-rays. Even if manual assessments of disc height, angles
13 and SPO had been obtained, error due to observer variability would be expected.^{11,21-23} Thus, analysis
14 without a gold standard was required. This prevented traditional sensitivity/specificity analysis.

15 Expert consensus is that a standing neutral-lateral radiograph should be used for diagnosis of
16 sagittal plane spondylolisthesis, although it can also be measured in flexion and extension.^{24,25} ²⁶ There is
17 limited published evidence about how flexion and extension may affect the assessment of
18 spondylolisthesis. Considering the Pearson correlation coefficients in Table 2, all three SPO metrics were
19 significantly correlated with disc angle. A variable amount of disc height loss can also coexist with
20 spondylolisthesis and there is limited evidence how variability in disc heights effects the assessment of
21 spondylolisthesis. The data from the NHANES-II study, as well as the analysis reported in Supplemental
22 Appendix 2 may help to establish comprehensive and standardized SPO assessment.

23 One of the most compelling pieces of evidence in support of the need for considering disc heights
24 and angles when assessing SPO was observed when applying the NHANES-II SPO reference data to
25 flexion-extension X-rays (Supplemental Appendix 2). Very large differences in the NHANES-II based

1 standardized ASPO, PSPO, and CSPO metrics were observed between the flexion and extension X-rays
2 for each subject. The NHANES-II standardized ASPO, PSPO, and CSPO metrics account for the
3 relatively low variability that occurs in a collection of X-rays that were all obtained in the lateral
4 decubitus position. The NHNAES-II SPO data are clearly not applicable to X-rays taken in a flexed or
5 extended position, unless the thresholds used to classify a level as normal vs abnormal are set to a high
6 level. In Supplemental Appendix 2 a “Spondylolisthesis Index” (SI) is developed that accounts for
7 variability in disc heights and disc angles. SI is also reported in units of deviation from average normal at
8 radiographically normal levels in asymptomatic volunteers.

9 The experiment using simulated X-rays, described in Supplemental Appendix 3, documents that
10 some error in the disc height, angle and SPO measurements occurs due to variability in radiographic
11 projections. As an example, the median error in measuring ADH from a radiograph was 0.01 std dev from
12 average normal ADH. This is the typical variability expected only due to radiographic projection.
13 However, for highly out-of-plane imaging of a disc space, errors can approach 1 std dev solely due to the
14 poorly imaged level. Solutions to this issue with lateral spine radiographs include protocols to optimize
15 imaging of each level and possibly neural networks trained to correct for out-of-plane imaging.

16 There are multiple approaches to both imaging and measuring spondylolisthesis and this must be
17 considered when attempting to compare data between studies.^{4,27} There is as yet no validated consensus
18 for how best to measure and interpret spondylolisthesis. The criteria for interpreting a spondylolisthesis
19 measurement must be specific to how the measurement is made to avoid incorrect classification of
20 spondylolisthesis as normal versus abnormal.²⁷ Reliability of the measurements depends on the method
21 used.^{23,27} Measurement error has been documented.²⁸⁻³³ In one study, spondylolisthesis was measured
22 with interclass correlation coefficients that were generally ≥ 0.8 , with better results when measured on a
23 workstation allowing zooming in on the anatomy.³⁴ Interclass correlation coefficients and other metrics
24 can document reproducibility, but accuracy is also important. Finally, optimizing diagnosis of

1 spondylolisthesis is dependent on understanding the variables that affect interpretation of the
2 measurements.

3 Clinical studies of treatments for symptomatic spondylolisthesis will typically report that patients
4 had spondylolisthesis but do not give any details about how the assessment was made, other than perhaps
5 a threshold amount of SPO (eg Austevoll et al³⁵). A threshold level of 3 mm is commonly used. However,
6 it is necessary to correct for image magnification when measuring SPO in millimeters, since
7 magnification can be highly variable.¹⁶ PSPO can normally vary by over 8.56 % endplate width based on
8 trimmed data in Supplemental Appendix 1 for the L4L5 level. Using 35 mm as a typical endplate width,
9 then approximately 3 mm of PSPO can be within normal limits. With high magnification, 3 mm of SPO
10 would likely not be abnormal. To avoid errors when radiographic magnification is unknown, it may be
11 preferable to measure SPO as % endplate width and use appropriate reference data.

12 The SPO measurement method that will prove of greatest clinical benefit has yet to be determined.
13 Sagittal plane displacement of the centroid of the vertebral body (Figure 3) may be important when the
14 displacement of the entire superior vertebral body relative to the inferior vertebra is prognostic. A more
15 localized measurement, such as the posterior-inferior corner of the superior vertebra relative to posterior-
16 superior corner of the inferior vertebra (PSPO) might be of greater significance when inflammation of the
17 tissue within the neural foramina immediately adjacent to the vertebral corner is a source of symptoms. The
18 descriptive statistics in Supplemental Appendix 1 document that, excluding outliers, ASPO, PSPO, or
19 CSPO is level dependent and rarely exceeds 14% endplate width.

20 In addition to measurement of the magnitude of SPO, there are different aspects of
21 spondylolisthesis that may be clinically important^{36,37}. The degenerative form of spondylolisthesis is
22 associated with arthritis and hypermobility of the facet joints.³⁸⁻⁴⁰ It has also been shown that hypoplasia of
23 L5 can occur and this can create a false impression of spondylolisthesis.⁴¹⁻⁴³ Vertebral morphology data
24 from NHANES-II can help to identify hypoplasia and other abnormalities. {Hipp, 2022 #10614} It may also
25 be important to interpret SPO with respect to factors such as pelvic incidence.^{31,44} An alternative

1 classification system was described by Marchetti-Bartolozzi.⁴⁵ The relative importance of SPO magnitude
2 versus the other aspects of classifying spondylolisthesis remains to be determined. It is likely that with
3 sufficient training material, artificial intelligence can be trained to objectively apply different
4 spondylolisthesis classification systems that are currently based primarily on subjective assessment. That
5 could supplement basic quantitative metrics.

6 Multiple prior studies help with assessing the prevalence of abnormal SPO in the NHANES-II as
7 documented in Supplemental Appendix 1. Kalichman et al found a prevalence of 21% in the Framingham
8 heart study, with the prevalence increasing with age.⁴⁶ That is higher than in the NHANES-II study, but
9 the population sampling is different. There is some evidence of an association with occupational
10 exposure.⁴⁷ There are relevant data collected in the NHANES-II study that could be use to test this and
11 similar associations. Importantly with respect to the spondylolisthesis index, Cehn et al found that the disc
12 height is a predictor of spondylolisthesis, as well as the ratio of posterior to anterior vertebral body
13 height.⁴⁸ It would be interesting to know what might be learned if all these studies were repeated with
14 objective and standardized SPO metrics.

15 Although SPO was a primary focus of the study, a lot of disc height and disc angle data were
16 generated that may prove useful. The disc height data in particular may help with assessing for loss in
17 disc height, which may prove important in multiple scenarios, such as identifying candidates for biologic
18 disc treatments where the treatment may be more effective in treating the early stages of disc
19 degeneration. The association of disc height with age in the NHANES-II study was significant, though
20 weak. Age explained < 1% of the variation in ADH and VDH, and 4% of the variation in PDH and DDH
21 ($P < 0.0001$). Whereas several authors have reported an association of disc height with age (either disc
22 height increases with age or decreases with age, depending on the study), the strength of the association is
23 very important to consider.⁴⁹⁻⁵⁴ It is also possible that reporting disc height in units of percent endplate
24 width minimize the differences between sexes compared to studies showing a difference in disc height
25 between males and females when disc height is measured in millimeters.^{53,54}

1 Despite that publications and spine fusion coverage policies infer interchangeability of
2 spondylolisthesis with instability⁵⁵, this has been proven in several studies NOT to be true.⁵⁶⁻⁵⁹ There is no
3 way to address this issue with the NHANES-II data, in large part due to the lack of validated diagnostic
4 metrics for instability.⁶⁰⁻⁶² Critical assessment of the available evidence base supports that the clinical
5 importance of, and optimal treatment for spondylolisthesis is not fully understood.⁶³⁻⁶⁶ Assessing whether
6 the spondylolisthesis is static or dynamic is likely very important^{67,68}. That could not be assessed using the
7 NHANES-II data. Development of the evidence base will likely be facilitated by objective and standardized
8 measurements that are consistently used across research studies.

9 Limitations of this study include:

- 10 1) Poor representation of some races, nations of origin, and sex in some age groups. In particular,
11 females are under-represented. By design of the NHANES-II study, no lumbar X-rays were taken
12 for pregnant women or women under 50.⁶⁹ An additional limitation relates to the data on race and
13 nation of origin recorded in the NHANES-II data. 86.9% of NHANES-II subjects were “white”,
14 11.2% “black” and the rest “other”. A more uniform representation of races would likely be
15 needed to fully understand the importance of race. The same is true for the “nation of origin” data
16 in NHANES-II. *NOTE: the language in quotes is as used in the NHANES-II documentation:*
17 74.1% of subjects are classified as “OTHER EUROPE SUCH AS GERMAN, FRENCH,
18 ENGLISH, IRISH”, 10.8% classified as “BLACK, NEGRO OR AFRO-AMERICAN”, and the
19 rest distributed in 1 of 11 other classifications. It would be valuable to repeat this analysis on X-
20 rays that better represent races and nations of origin that are not well represented in the
21 NHANES-II data. The ages of the subjects were also biased toward older ages (Figure 4).
- 22 2) NHANES-II lumbar spine X-rays were obtained side-lying. Disc metrics in Appendix 1 may not
23 be directly and precisely applicable to X-rays taken in other positions. Using all of the NHANES
24 data, the L1-S1 angle was 51.4 (std Dev 12.7). Table 4 provides data from three prior studies
25 reporting L1-S1 angles measured from X-rays obtained with subjects standing.⁷⁰⁻⁷² The lordosis

1 data from the NHANES-II study is very similar to lordosis measured from standing X-rays in
2 three other studies. Figure 5 shows mid-plane angles from the NHANES-II study, excluding
3 levels where any disc or SPO metric was abnormal, compared to mid-plane angles reported by
4 Frobin et al from X-rays obtained with subjects standing.¹⁴ There are some differences, but also
5 some similarities. These comparative data can help in deciding whether to use NHANES-II
6 reference data with standing or other X-rays. With respect to external validity of the disc height
7 measurements, comparable reference data were hard to find. There are multiple publications
8 reporting lumbar intervertebral disc height reference data in units of millimeters, frequently from
9 MRI or CT exams. Since a scaling device was not included in the NHANES-II X-rays, these prior
10 publications cannot be used as comparative data. Data for a large collection of radiographs
11 obtained in other positions and analyzed using the same methods would be required to understand
12 applicability of the NHANES-II data to other protocols for obtaining lateral lumbar spine X-rays.
13 It is likely that the average disc height data will be the most universally applicable of the
14 NHANES-II reference data, since average disc height should minimize minor differences due to
15 positioning.

16 3) The accuracy study using simulated X-rays (Supplemental Appendix 3) revealed potentially large
17 errors in disc and SPO metrics when the radiographic projection is poor. It would be valuable to
18 establish a neural network or other method to either make a correction in the metrics when large
19 out-of-plane imaging is encountered, or to abstain from reporting data when vertebrae are poorly
20 imaged. Nevertheless, a certain level of error must be expected in SPO and other metrics. This
21 may be a particularly significant issue with large amounts of frontal plane spinal curvature. This
22 study does not provide clear guidance on this issue and additional data are needed. Caution
23 should be used when interpreting disc metrics and SPO from lateral X-rays where the vertebrae
24 are poorly imaged.

25 4) One of the greatest challenges was in trimming the data such that only truly normal discs were
26 used to define “normal” disc and SPO metrics. Abnormal disc heights, angles, and SPO can be

1 expected in the NHANES-II study (since there was no attempt to exclude spine abnormalities). It
2 was assumed that normal disc heights, disc angles, and SPO would have a gaussian distribution.
3 It was also assumed that discs with abnormal heights, angles, or SPO, due to degeneration,
4 trauma, genetics or other reasons would be in the tails of the distributions when all NHANES-II
5 X-rays are combined. Even in healthy discs, a range of disc heights, angles, and SPO would be
6 expected due to genetic differences between individuals (and possibly nutritional history and
7 other variables). In addition, in a normal disc, the resting position of the vertebrae, when the
8 radiograph was obtained, would be expected to be in the neutral zone. Within the neutral zone,
9 little force is required to produce sagittal plane movements.^{73,74} It can be hypothesized that the
10 amount of SPO may change slightly, but remain within the neutral zone, every time a person
11 assumes a “neutral” position. That would be expected to contribute to a normal distribution of
12 SPO within a population. There is currently no gold standard method that is validated for
13 classifying discs as normal versus abnormal. MRI exams were not available. In the current study,
14 an attempt was made to trim data so as to achieve a gaussian distribution. This approach has been
15 used by other authors.⁷⁵⁻⁷⁷ A better strategy may be possible for assuring that only truly normal
16 discs are used to define normative reference data. For example, a formal optimization scheme
17 might be used to obtain a gaussian of a distribution, though justification for such an optimization
18 is not well developed.

19 5) Although Supplemental Appendix 1 provides data to calculate disc and SPO metrics as number of
20 Std Dev from average normal, a clinically meaningful threshold must be validated for use with
21 these reference data that can classify a metric as normal vs abnormal. Even though a metric that is
22 two Std Dev from average would be technically outside of the 95% confidence interval used to
23 define “normal”, that may not be clinically significant. Until well-designed clinical trials are
24 completed, the threshold level for the standardized score that is clinically efficacious will not be
25 known.

- 1 Table 4: L1-S1 angles from three prior studies⁷⁰⁻⁷² measured from standing lateral x-rays compared to L1-
2 S1 angles in the NHANES-II study

	N	Mean Age	Mean L1-S1	StdDev L1-S1
Korovessis ⁷⁰	45	63	48	12.5
Lord ⁷¹	109	47	49	15
Been ⁷²	100	38	51	
NHANES-II	6397	55	51	13

3
4 From a practical standpoint, degenerative spondylolisthesis has traditionally been considered a
5 single diagnosis. It is now clear that this in fact is a very heterogenous condition that requires a deeper
6 understanding to effectively optimize its management. Beyond simply recognizing that this
7 radiographic interpretation has a depth of complexity, we are also limited by our relatively primitive
8 and inaccurate assessment tools. When a patient record or a scientific paper reports that a patient or
9 patients had disc height loss or spondylolisthesis, what evidence exists that the report is accurate?
10 What evidence exists that the magnitude of SPO is reliably reported? Will disc height and
11 spondylolisthesis measurements prove valuable in diagnosis and treatment? It seems probable that
12 clinical trials (past and future) might yield better and more actionable knowledge if accurate,
13 reproducible, and standardized assessments are used. From a clinician's perspective, there is
14 something about spondylolisthesis that is contributing to clinical impairment in some but not all
15 cases. Understanding the clinical predictors that forecast the role of spondylolisthesis in symptom
16 generation both before and after surgery will be critical to optimal management of this prevalent
17 condition. With developing artificial intelligence, it appears we are much closer to objectively and
18 accurately evaluating and quantifying degenerative spondylolisthesis and automatically obtaining
19 these metrics, with minimal effort, in routine practice. The NHANES-II reference data and
20 measurements must be critically correlated with clinical metrics including but not limited to
21 subjective complaints, physical exam, and sociodemographic features of the patient. Artificial
22 intelligence will likely automate routine and basic measurements such as disc height loss and

1 spondylolisthesis, such that the clinician can focus on integrating the assessments. Integrating
2 imaging with clinical characteristics should allow development of reliable predictive analytics to
3 optimize clinical and surgical decision making in the future.

4 Of note: there would be minimal technical challenges or cost to re-analysis of prior studies using
5 objective SPO metrics to determine if the amount of SPO can help to predict treatment outcomes.

6 **Conclusion**

7
8 The NHANES-II radiographs were used to develop reference data to aid in interpreting
9 measurements of sagittal plane offset (SPO), disc heights, and disc angles. These metrics are dependent
10 on disc level, so predicting normal SPO should always be done with level-specific reference data. The
11 NHANES-II data document that SPO is significantly dependent on disc heights and disc angles, though
12 the effect is small when all X-rays are obtained in a neutral position. Analysis of standing flexion-
13 extension radiographs documents a far more significant effect of disc height and disc angle on SPO, and
14 thus the importance of evaluating disc height and angle when predicting what should be normal SPO. The
15 data support predicting normal SPO based only on level as long as there is little variability between
16 subjects independent of the position of the spine, but disc heights and angles should be used to predict
17 normal SPO when the spine is flexed or extended. The research and clinical value of any quantitative
18 diagnostic is improved when sources of measurement error are minimized. Reference data developed
19 from the NHANES-II lumbar X-rays provide normative reference data for predicting normal SPO, disc
20 height, and disc angles. These reference data would need to be tested for efficacy in appropriate clinical
21 trials. Multiple metrics were assessed using the NHANES-II data, yet it will take further research to
22 determine which are clinically efficacious. The greatest challenge is to validate diagnosis and treatment
23 algorithms that can effectively make use of disc and SPO metrics to improve outcomes for patients with
24 spinal disorders.

25

1 **Ethics Approval**

2

3 Vertebral morphology was measured from lateral lumbar and cervical radiographs obtained during the 2nd

4 National Health and Nutrition Survey (1976-1980). That study was designed and conducted to protect the

5 rights of the volunteers, and all subjects in the study signed a consent as described in the study

6 documentation:

7 <https://www.cdc.gov/nchs/nhanes/nhanes2/manualsandreports.aspx>

8 https://www.cdc.gov/nchs/data/nhanes2/manuals/15a76_79.pdf

9

10

1 Figure Legends

2 Figure 1: Disc height and SPO measurements using method 1.

3 Figure 2: Disc height and SPO measurements using method 2 (adapted from Frobin et al¹⁴).

4 Figure 3: Method used to measure Centroid Sagittal Plane Offset (CSPO), adapted from Frobin et al.¹⁵.

5 The dotted lines connect the centroids to the disc bisectrix and are perpendicular to the bisectrix. CSPO
6 therefore measures the offset between the vertebral centroids along the disc bisectrix.

7 Figure 4: Distribution of ages in the NHANES-II study.

8 Figure 5: A comparison of mid-plane angles from the NHANES-II study (excluding all levels where any
9 metric was abnormal) to data from Frobin et al.¹⁴ X-rays in the NHANES-II study were obtained with
10 subjects lying on their sides in a semi flexed position, while Frobin et al obtained X-rays with patients
11 standing. The error bars show standard deviations. The sample size in Frobin et al was 61 to 239
12 depending on sex and level, while the sample size in the NHANES-II data was 1185 to 2572 depending
13 on sex and level.

14

15

16

17

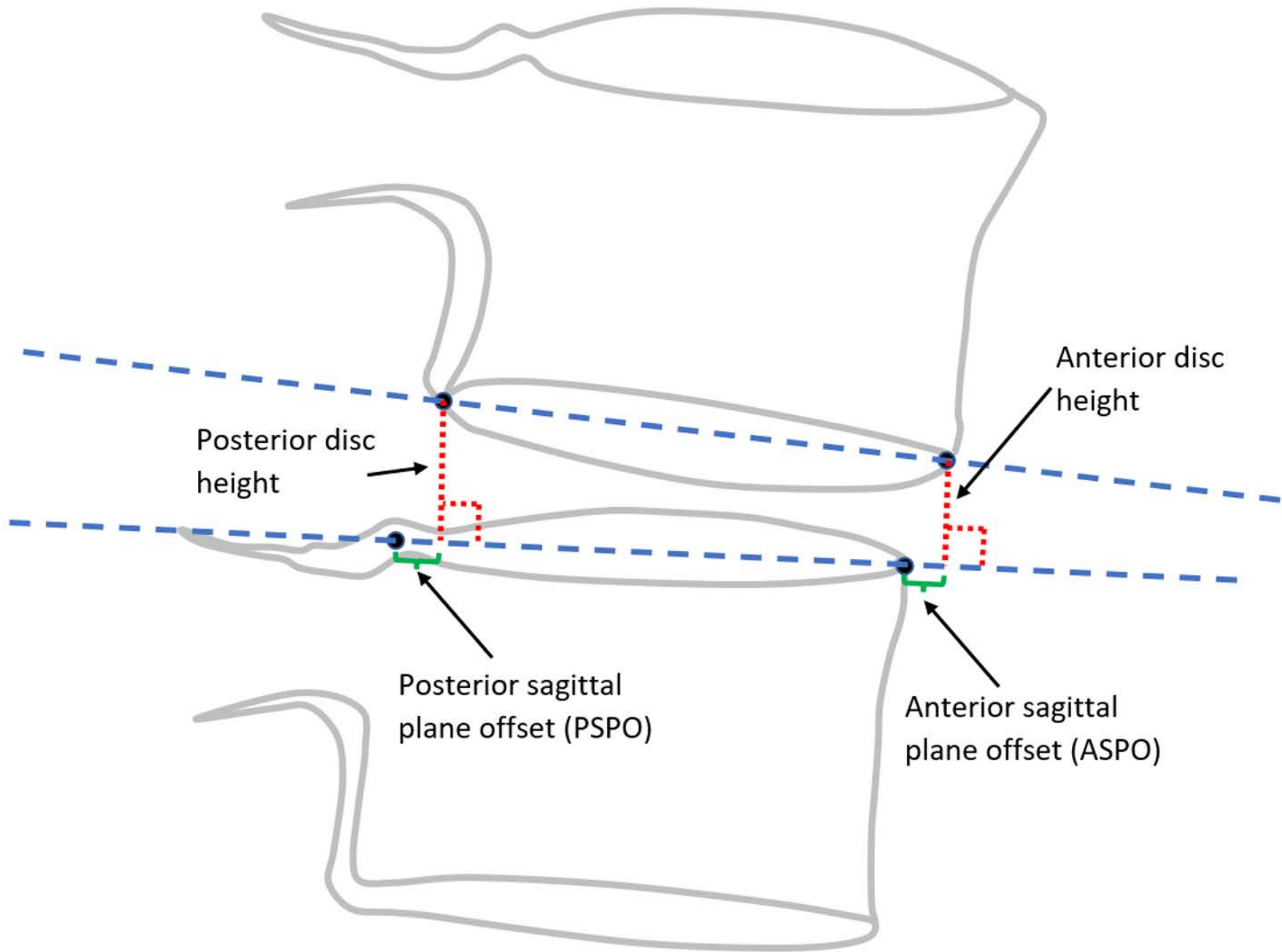
1 References

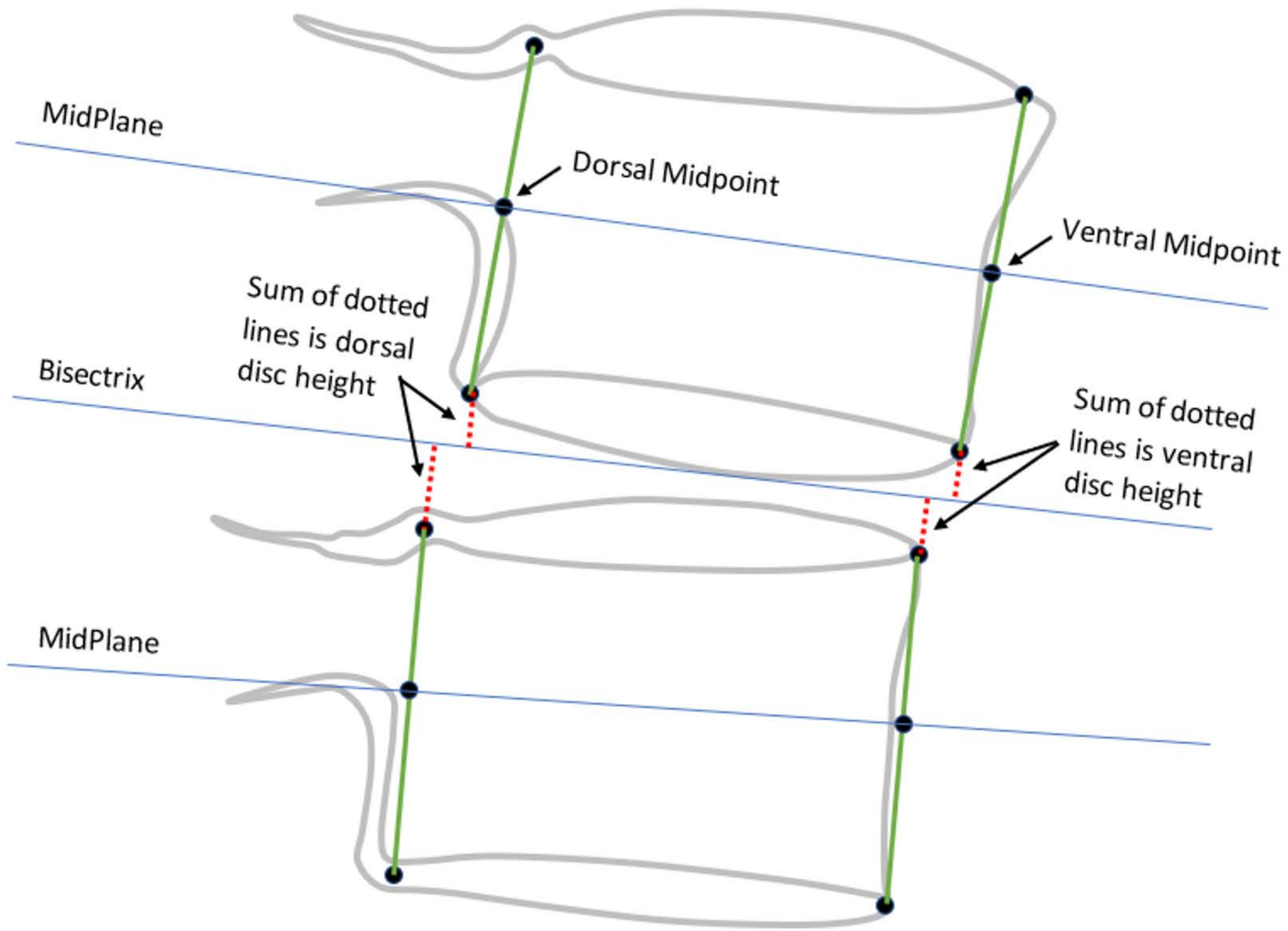
- 2 1. Galbusera F, Casaroli G, Bassani T. Artificial intelligence and machine learning in spine
3 research. *Jor Spine*. 2019;2(1):e1044.
- 4 2. Chang M, Canseco JA, Nicholson KJ, Patel N, Vaccaro AR. The role of machine learning in
5 spine surgery: the future is now. *Frontiers in surgery*. 2020;7:54.
- 6 3. Ren G, Yu K, Xie Z, et al. Current Applications of Machine Learning in Spine: From Clinical
7 View. *Global Spine Journal*. 2021:21925682211035363.
- 8 4. Câmara JR, Keen JR, Asgarzadie F. Functional radiography in examination of spondylolisthesis.
9 *American Journal of Roentgenology*. 2015;204(4):W461-W469.
- 10 5. McDowell A. Plan and operation of the second National Health and Nutrition Examination
11 Survey, 1976-1980. 1981;
- 12 6. Zhao KD, Ben-Abraham EI, Magnuson DJ, et al. Effect of Off-Axis Fluoroscopy Imaging on
13 Two-Dimensional Kinematics in the Lumbar Spine: A Dynamic In Vitro Validation Study. *J Biomech*
14 *Eng*. May 2016;138(5):054502. doi:10.1115/1.4032995
- 15 7. Zhao KD, Yang C, Zhao C, Stans AA, An KN. Assessment of noninvasive intervertebral motion
16 measurements in the lumbar spine. *JBiomechanics*. 2005;38(9):1943-1946. NOT IN FILE.
17 doi:10.1016/j.jbiomech.2004.07.029
- 18 8. Reitman CA, Hipp JA, Nguyen L, Esses SI. Changes in segmental intervertebral motion adjacent
19 to cervical arthrodesis: a prospective study. *Spine (Phila Pa 1976)*. Jun 1 2004;29(11):E221-6.
20 doi:00007632-200406010-00022 [pii]
- 21 9. Reitman CA, Mauro KM, Nguyen L, Ziegler JM, Hipp JA. Intervertebral motion between flexion
22 and extension in asymptomatic individuals. *Spine (Phila Pa 1976)*. Dec 15 2004;29(24):2832-43.
23 doi:00007632-200412150-00009 [pii]
- 24 10. Patwardhan AG, Havey RM, Wharton ND, et al. Asymmetric motion distribution between
25 components of a mobile-core lumbar disc prosthesis: an explanation of unequal wear distribution in
26 explanted CHARITE polyethylene cores. *The Journal of bone and joint surgery American volume*. May 2
27 2012;94(9):846-54. doi:10.2106/JBJS.J.00638
- 28 11. Pearson AM, Spratt KF, Genuario J, et al. Precision of lumbar intervertebral measurements: Does
29 a computer-assisted technique improve reliability? *Spine*. 2011;36(7):572-580.
30 doi:10.1097/BRS.0b013e3181e11c13
- 31 12. Pfeiffer M, Geisel T. Analysis of a computer-assisted technique for measuring the lumbar spine
32 on radiographs: correlation of two methods. *Academic radiology*. 2003;10(3):275-282.
- 33 13. Hipp JA, Grieco TF, Newman P, Reitman CA. Definition of normal vertebral morphology using
34 NHANES-II radiographs. *medRxiv*. 2022;
- 35 14. Frobin W, Brinckmann P, Biggemann M, Tillotson M, Burton K. Precision measurement of disc
36 height, vertebral height and sagittal plane displacement from lateral radiographic views of the lumbar
37 spine. *Clin Biomech (Bristol, Avon)*. 1997;12 Suppl 1:S1-S63. doi:10.1016/s0268-0033(96)00067-8
- 38 15. Frobin W, Brinckmann P, Leivseth G, Biggemann M, Reikeras O. Precision measurement of
39 segmental motion from flexion-extension radiographs of the lumbar spine. *ClinBiomech(Bristol, Avon)*.
40 1996;11(8):457-465. NOT IN FILE.
- 41 16. Ravi B, Rampersaud R. Clinical magnification error in lateral spinal digital radiographs. *Spine*.
42 2008;33(10):E311-E316.
- 43 17. Shigematsu H, Koizumi M, Yoneda M, Iida J, Oshima T, Tanaka Y. Magnification error in
44 digital radiographs of the cervical spine against magnetic resonance imaging measurements. *Asian spine*
45 *journal*. 2013;7(4):267.
- 46 18. Lunt M, Gowin W, Johnell, Armbrrecht G, Felsenberg D. A statistical method to minimize
47 magnification errors in serial vertebral radiographs. *OsteoporosInt*. 2001;12(11):909-913. NOT IN FILE.
- 48 19. Staub BN, Holman PJ, Reitman CA, Hipp JA. Sagittal plane lumbar intervertebral motion during
49 seated flexion-extension in 658 asymptomatic, non-degenerated levels. *J Neurosurgery Spine*. 2015
50 2015;23(6):731-738. doi:10.3171/2015.3.SPINE14898

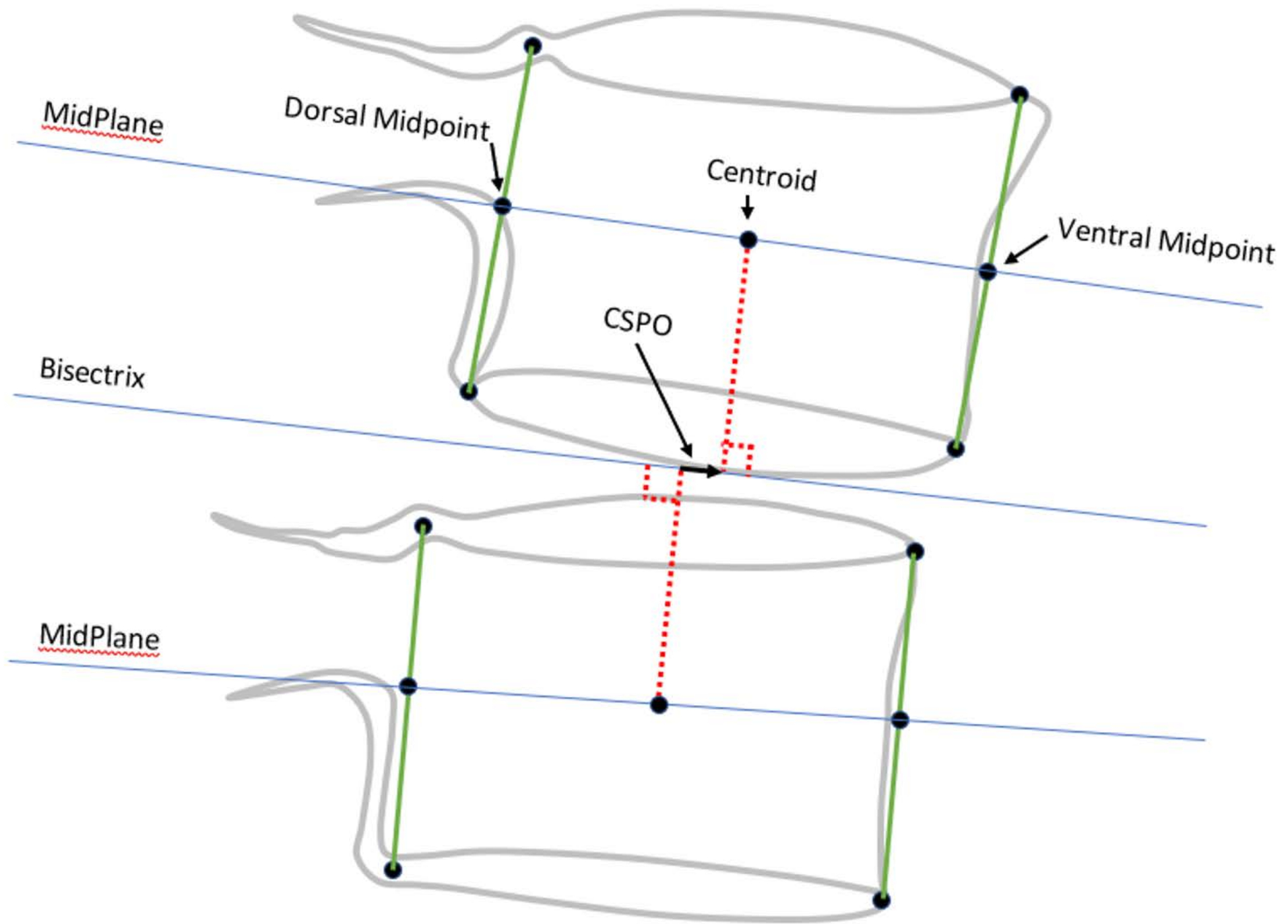
- 1 20. Bifulco P, Sansone M, Cesarelli M, Allen R, Bracale M. Estimation of out-of-plane vertebra
2 rotations on radiographic projections using CT data: a simulation study. *MedEng Phys.* 2002;24(4):295-
3 300. NOT IN FILE.
- 4 21. Andersson G, Schultz A, Nathan A, Irstam L. Roentgenographic measurement of lumbar
5 intervertebral disc height. *Spine.* 1981;6(2):154-158.
- 6 22. Danielson B, Frennered K, Irstam L. Roentgenologic assessment of spondylolisthesis: I. A study
7 of measurement variations. *Acta Radiologica.* 1988;29(3):345-351.
- 8 23. McCarty ME, Mehlman CT, Tamai J, Do TT, Crawford AH, Klein G. Spondylolisthesis:
9 intraobserver and interobserver reliability with regard to the measurement of slip percentage. *Journal of*
10 *Pediatric Orthopaedics.* 2009;29(7):755-759.
- 11 24. Watters WC, 3rd, Bono CM, Gilbert TJ, et al. An evidence-based clinical guideline for the
12 diagnosis and treatment of degenerative lumbar spondylolisthesis. *Spine J.* Jul 2009;9(7):609-14.
13 doi:10.1016/j.spinee.2009.03.016
- 14 25. Matz PG, Meagher R, Lamer T, et al. Guideline summary review: an evidence-based clinical
15 guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. *The Spine Journal.*
16 2016;16(3):439-448.
- 17 26. Mummaneni PV, Bisson EF, Kerezoudis P, et al. Minimally invasive versus open fusion for
18 Grade I degenerative lumbar spondylolisthesis: analysis of the Quality Outcomes Database.
19 *Neurosurgical focus.* 2017;43(2):E11.
- 20 27. Bourassa-Moreau É, Mac-Thiong J-M, Labelle H. Redefining the technique for the radiologic
21 measurement of slip in spondylolisthesis. *Spine.* 2010;35(14):1401-1405.
- 22 28. Danielson B, Frennered K, Selvik G, Irstam L. Roentgenologic Assessment of Spondylolisthesis:
23 II. An Evaluation of Progression. *Acta Radiologica.* 1989;30(1):65-68.
- 24 29. Wall MS, Oppenheim WL. Measurement error of spondylolisthesis as a function of radiographic
25 beam angle. *Journal of pediatric orthopedics.* 1995;15(2):193-198.
- 26 30. Tallroth K, Ylikoski M, Landtman M, Santavirta S. Reliability of radiographical measurements of
27 spondylolisthesis and extension-flexion radiographs of the lumbar spine. *Eur J Radiol.* 1994;18(3):227-
28 231. NOT IN FILE.
- 29 31. Mac-Thiong JM, Duong L, Parent S, et al. Reliability of the Spinal Deformity Study Group
30 classification of lumbosacral spondylolisthesis. *Spine (Phila Pa 1976).* Jan 15 2012;37(2):E95-102.
31 doi:10.1097/BRS.0b013e3182233969
- 32 32. Bao H, Yan P, Zhu W, et al. Validation and reliability analysis of the spinal deformity study
33 group classification for L5-S1 lumbar spondylolisthesis. *Spine.* 2015;40(21):E1150-E1154.
- 34 33. Denard PJ, Holton KF, Miller J, et al. Lumbar spondylolisthesis among elderly men: prevalence,
35 correlates and progression. *Spine.* 2010;35(10):1072.
- 36 34. Bolesta MJ, Winslow L, Gill K. A comparison of film and computer workstation measurements
37 of degenerative spondylolisthesis: intraobserver and interobserver reliability. *Spine.* 2010;35(13):1300-
38 1303.
- 39 35. Austevoll IM, Gjestad R, Brox JI, et al. The effectiveness of decompression alone compared with
40 additional fusion for lumbar spinal stenosis with degenerative spondylolisthesis: a pragmatic comparative
41 non-inferiority observational study from the Norwegian Registry for Spine Surgery. *European Spine*
42 *Journal.* 2017;26(2):404-413.
- 43 36. Wiltse LL, Newman PH, Macnab I. Classification of spondylolysis and spondylolisthesis. *Clin*
44 *Orthop Relat Res.* 1976;(117):23-29. NOT IN FILE.
- 45 37. Wiltse LL. Classification, terminology and measurements in spondylolisthesis. *The Iowa*
46 *orthopaedic journal.* 1981;1:52.
- 47 38. Newman PH, Stone KH. The etiology of spondylolisthesis. *Journal of Bone & Joint Surgery,*
48 *British Volume.* 1963;45(1):39-59. NOT IN FILE.
- 49 39. Fujiwara A, Tamai K, An HS, et al. The relationship between disc degeneration, facet joint
50 osteoarthritis, and stability of the degenerative lumbar spine. *JSpinal Disord.* 2000;13(5):444-450. NOT
51 IN FILE.

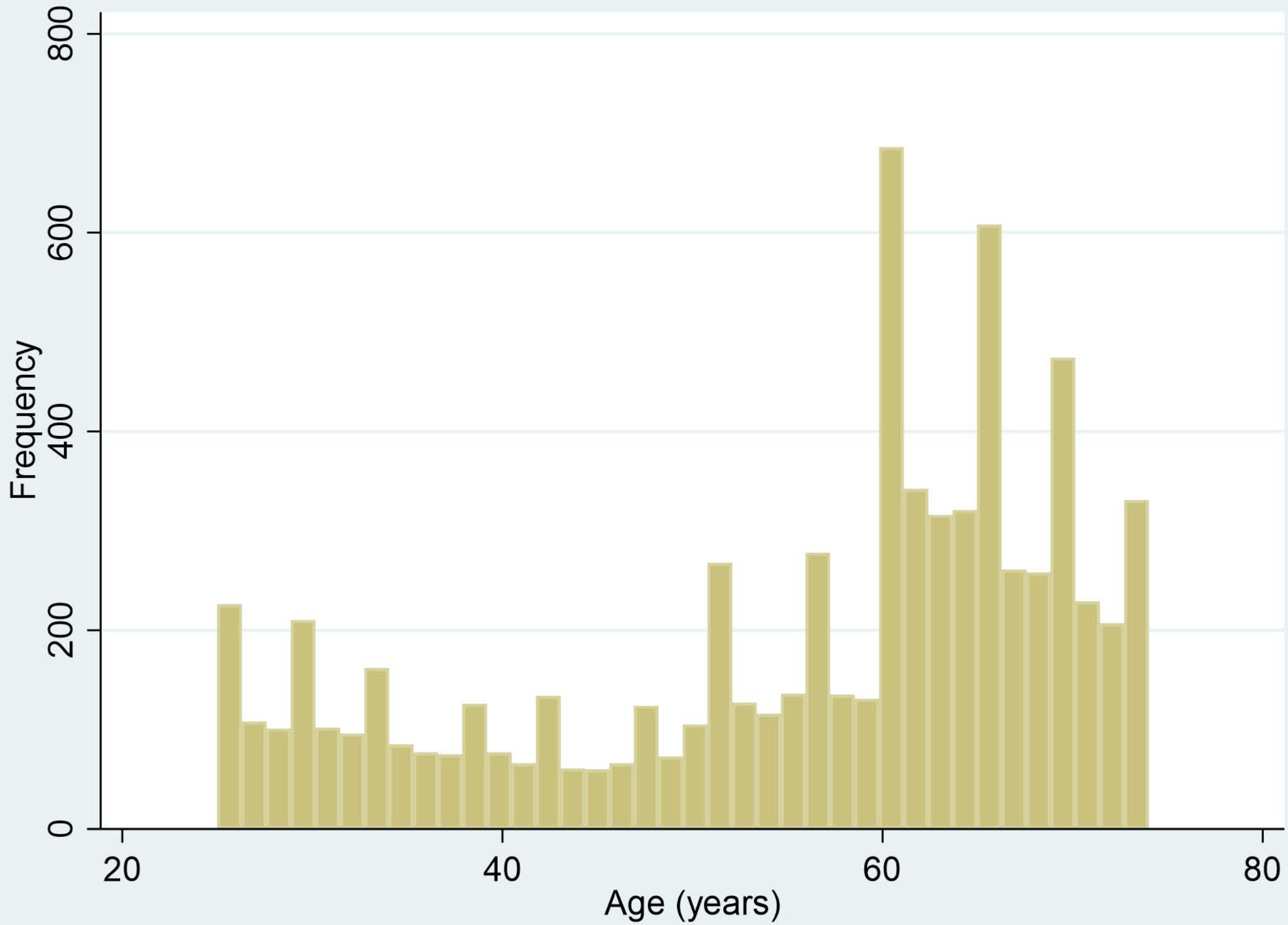
- 1 40. Chaput C, Padon D, Rush J, Lenahan E, Rahm M. The significance of increased fluid signal on
2 magnetic resonance imaging in lumbar facets in relationship to degenerative spondylolisthesis. *Spine*.
3 2007;32(17):1883-1887. NOT IN FILE.
- 4 41. Frank DF, Miller JE. Hypoplasia of the lumbar vertebral body simulating spondylolisthesis.
5 *Radiology*. 1979;133(1):59-60.
- 6 42. Niggemann P, Kuchta J, Grosskurth D, Beyer H, Hoeffler J, Delank K. Spondylolysis and isthmic
7 spondylolisthesis: impact of vertebral hypoplasia on the use of the Meyerding classification. *The British*
8 *journal of radiology*. 2012;85(1012):358-362.
- 9 43. Legaye J. Radiographic analysis of the listhesis associated with lumbar isthmic spondylolysis.
10 *Orthopaedics & Traumatology: Surgery & Research*. 2018;104(5):569-573.
- 11 44. Mac-Thiong J-M, Labelle H, Parent S, Hresko MT, Deviren V, Weidenbaum M. Reliability and
12 development of a new classification of lumbosacral spondylolisthesis. *Scoliosis*. 2008;3(1):1-9.
- 13 45. Wang YXJ, Kaplar Z, Deng M, Leung JC. Lumbar degenerative spondylolisthesis epidemiology:
14 a systematic review with a focus on gender-specific and age-specific prevalence. *Journal of orthopaedic*
15 *translation*. 2017;11:39-52.
- 16 46. Kalichman L, Kim DH, Li L, Guermazi A, Berkin V, Hunter DJ. Spondylolysis and
17 spondylolisthesis: prevalence and association with low back pain in the adult community-based
18 population. *Spine*. 2009;34(2):199.
- 19 47. Ishimoto Y, Cooper C, Ntani G, et al. Is radiographic lumbar spondylolisthesis associated with
20 occupational exposures? Findings from a nested case control study within the Wakayama spine study.
21 *BMC Musculoskeletal Disorders*. 2019;20(1):618.
- 22 48. Chen I-R, Wei T-S. Disc height and lumbar index as independent predictors of degenerative
23 spondylolisthesis in middle-aged women with low back pain. *Spine*. 2009;34(13):1402-1409.
- 24 49. Shao Z, Rompe G, Schiltenswolf M. Radiographic changes in the lumbar intervertebral discs and
25 lumbar vertebrae with age. *Spine*. 2002;27(3):263-268. NOT IN FILE.
- 26 50. Amonoo-Kuofi HS. Morphometric changes in the heights and anteroposterior diameters of the
27 lumbar intervertebral discs with age. *Journal of Anatomy*. 1991;175:159. NOT IN FILE.
- 28 51. Twomey LT, Taylor JR. Age changes in lumbar vertebrae and intervertebral discs. *Clinical*
29 *Orthopaedics & Related Research*. Nov 1987;(224):97-104. NOT IN FILE.
- 30 52. Videman T, Battié MC, Gibbons LE, Gill K. Aging changes in lumbar discs and vertebrae and
31 their interaction: a 15-year follow-up study. *The Spine Journal*. 2014;14(3):469-478.
- 32 53. Bach K, Ford J, Foley R, et al. Morphometric analysis of lumbar intervertebral disc height: an
33 imaging study. *World neurosurgery*. 2019;124:e106-e118.
- 34 54. Demir M, Emre A, Seringeç N, et al. Intervertebral disc heights and concavity index of the
35 lumbar spine in young healthy adults. *Anatomy*. 2018;12(1):34-37.
- 36 55. Leone A, Cassar-Pullicino VN, Guglielmi G, Bonomo L. Degenerative lumbar intervertebral
37 instability: what is it and how does imaging contribute? *Skeletal radiology*. 2009;38(6):529-533.
- 38 56. Lattig F, Fekete TF, Grob D, Kleinstuck FS, Jeszenszky D, Mannion AF. Lumbar facet joint
39 effusion in MRI: a sign of instability in degenerative spondylolisthesis? *European spine journal : official*
40 *publication of the European Spine Society, the European Spinal Deformity Society, and the European*
41 *Section of the Cervical Spine Research Society*. Feb 2012;21(2):276-81. doi:10.1007/s00586-011-1993-1
- 42 57. Simmonds AM, Rampersaud YR, Dvorak MF, Dea N, Melnyk AD, Fisher CG. Defining the
43 inherent stability of degenerative spondylolisthesis: a systematic review. *Journal of Neurosurgery: Spine*.
44 2015:1-12.
- 45 58. Hasegawa K, Kitahara K, Shimoda H, et al. Lumbar degenerative spondylolisthesis is not always
46 unstable: clinicobiomechanical evidence. *Spine*. 2014;39(26):2127-2135.
- 47 59. Wang D, Yuan H, Liu A, et al. Analysis of the relationship between the facet fluid sign and
48 lumbar spine motion of degenerative spondylolytic segment using Kinematic MRI. *European journal of*
49 *radiology*. 2017;94:6-12.
- 50 60. Nizard RS, Wybier M, Laredo JD. Radiologic assessment of lumbar intervertebral instability and
51 degenerative spondylolisthesis. *The Radiologic clinics of North America*. 2001;39(1):55-71.

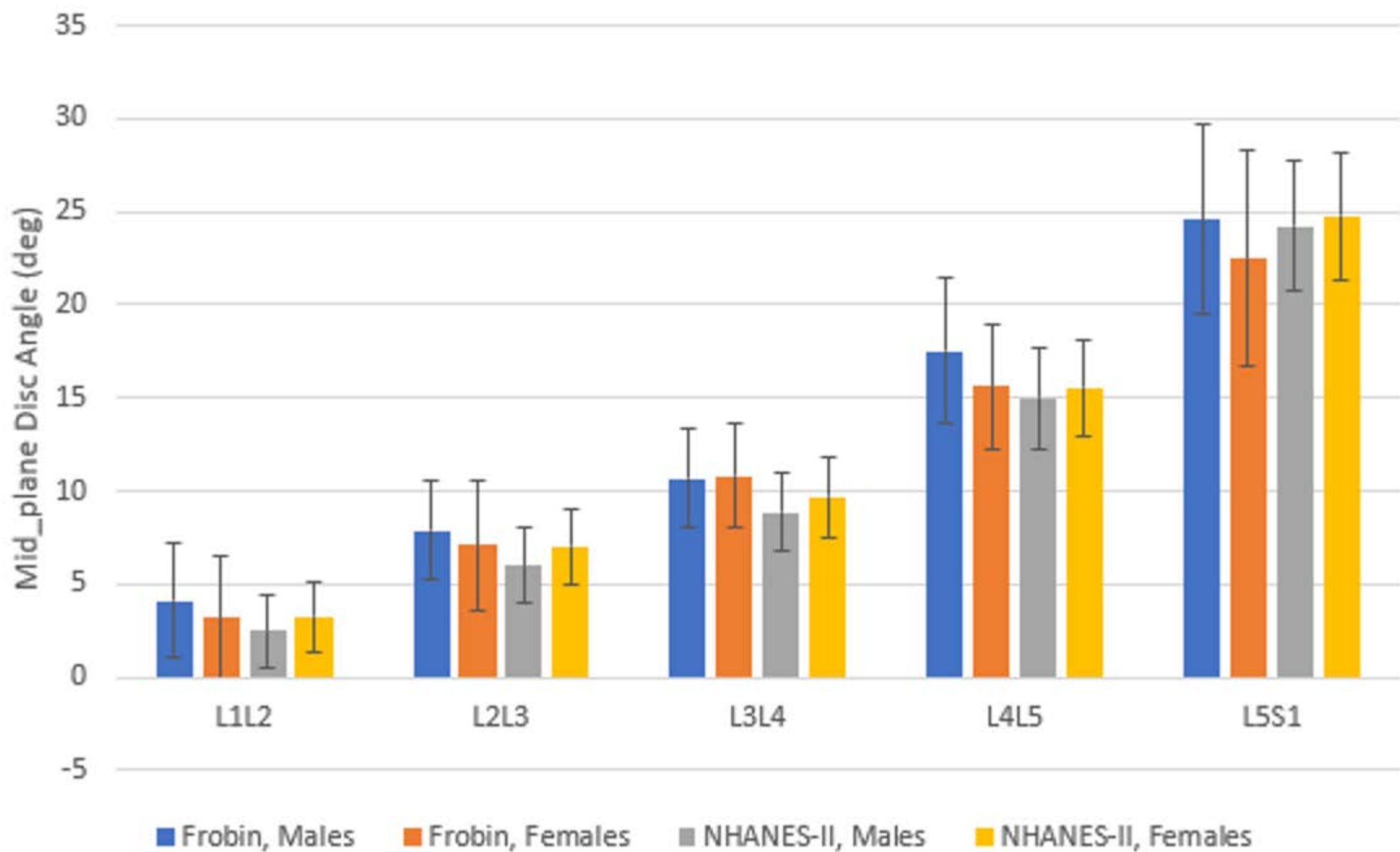
- 1 61. Leone A, Guglielmi G, Cassar-Pullicino VN, Bonomo L. Lumbar intervertebral instability: a
2 review. *Radiology*. Oct 2007;245(1):62-77. doi:10.1148/radiol.2451051359
- 3 62. Izzo R, Guarnieri G, Guglielmi G, Muto M. Biomechanics of the spine. Part I: spinal stability.
4 *Eur J Radiol*. Jan 2013;82(1):118-26. doi:10.1016/j.ejrad.2012.07.024
- 5 63. Evans N, McCarthy M. Management of symptomatic degenerative low-grade lumbar
6 spondylolisthesis. *EFORT open reviews*. 2018;3(12):620-631.
- 7 64. Wang M, Luo XJ, Ye YJ, Zhang Z. Does Concomitant Degenerative Spondylolisthesis Influence
8 the Outcome of Decompression Alone in Degenerative Lumbar Spinal Stenosis? A Meta-Analysis of
9 Comparative Studies. *World neurosurgery*. 2019;123:226-238.
- 10 65. Reitman CA, Cho CH, Bono CM, et al. Management of degenerative spondylolisthesis:
11 development of appropriate use criteria. *The Spine Journal*. 2021;21(8):1256-1267.
- 12 66. Mannion A, Pittet V, Steiger F, Vader J-P, Becker H-J, Porchet F. Development of
13 appropriateness criteria for the surgical treatment of symptomatic lumbar degenerative spondylolisthesis
14 (LDS). *European spine journal*. 2014;23(9):1903-1917.
- 15 67. Even JL, Chen AF, Lee JY. Imaging characteristics of “dynamic” versus “static”
16 spondylolisthesis: analysis using magnetic resonance imaging and flexion/extension films. *The Spine
17 Journal*. 2014;14(9):1965-1969.
- 18 68. Inose H, Kato T, Onuma H, et al. Predictive Factors Affecting Surgical Outcomes in Patients with
19 Degenerative Lumbar Spondylolisthesis. *Spine*. 2021;46(9):610-616.
- 20 69. Antani S, Cheng J, Long J, Long LR, Thoma GR. Medical validation and CBIR of spine x-ray
21 images over the Internet. *International Society for Optics and Photonics*; 2006:60610J.
- 22 70. Korovessis P, Repantis T, Papazisis Z, Iliopoulos P. Effect of Sagittal Spinal Balance, Levels of
23 Posterior Instrumentation, and Length of Follow-up on Low Back Pain in Patients Undergoing Posterior
24 Decompression and Instrumented Fusion for Degenerative Lumbar Spine Disease: A Multifactorial
25 Analysis. *Spine*. 2010;35(8):898.
- 26 71. Lord MJ, Small JM, Dinsay JM, Watkins RG. Lumbar lordosis. Effects of sitting and standing.
27 *Spine*. 1997;22(21):2571-2574. NOT IN FILE.
- 28 72. Been E, Barash A, Pessah H, Peleg S. A new look at the geometry of the lumbar spine. *Spine*.
29 2010;35(20):E1014-E1017.
- 30 73. Crawford NR, Peles JD, Dickman CA. The spinal lax zone and neutral zone: measurement
31 techniques and parameter comparisons. *J Spinal Disord*. Oct 1998;11(5):416-29.
- 32 74. Panjabi MM. The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis.
33 *J Spinal Disord*. Dec 1992;5(4):390-6; discussion 397. NOT IN FILE. doi:10.1097/00002517-
34 199212000-00002
- 35 75. Black DM, Cummings SR, Stone K, Hudes E, Palermo L, Steiger P. A new approach to defining
36 normal vertebral dimensions. *J Bone Min Res*. 1991;6:883-892. NOT IN FILE.
- 37 76. Rea J, Steiger P, Blake G, Potts E, Smith I, Fogelman I. Morphometric X-ray absorptiometry:
38 reference data for vertebral dimensions. *Journal of Bone and Mineral Research*. 1998;13(3):464-474.
- 39 77. O'Neill T, Varlow J, Felsenberg D, et al. Variation in vertebral height ratios in population studies.
40 *J Bone Miner Res*. 1994;12:1895-1907. NOT IN FILE.











Supplemental Appendix 1: Details of landmark placement, descriptive statistics, examples of abnormalities in the NHANES-II x-rays, and prevalence of abnormalities in the NHANES-II x-rays.

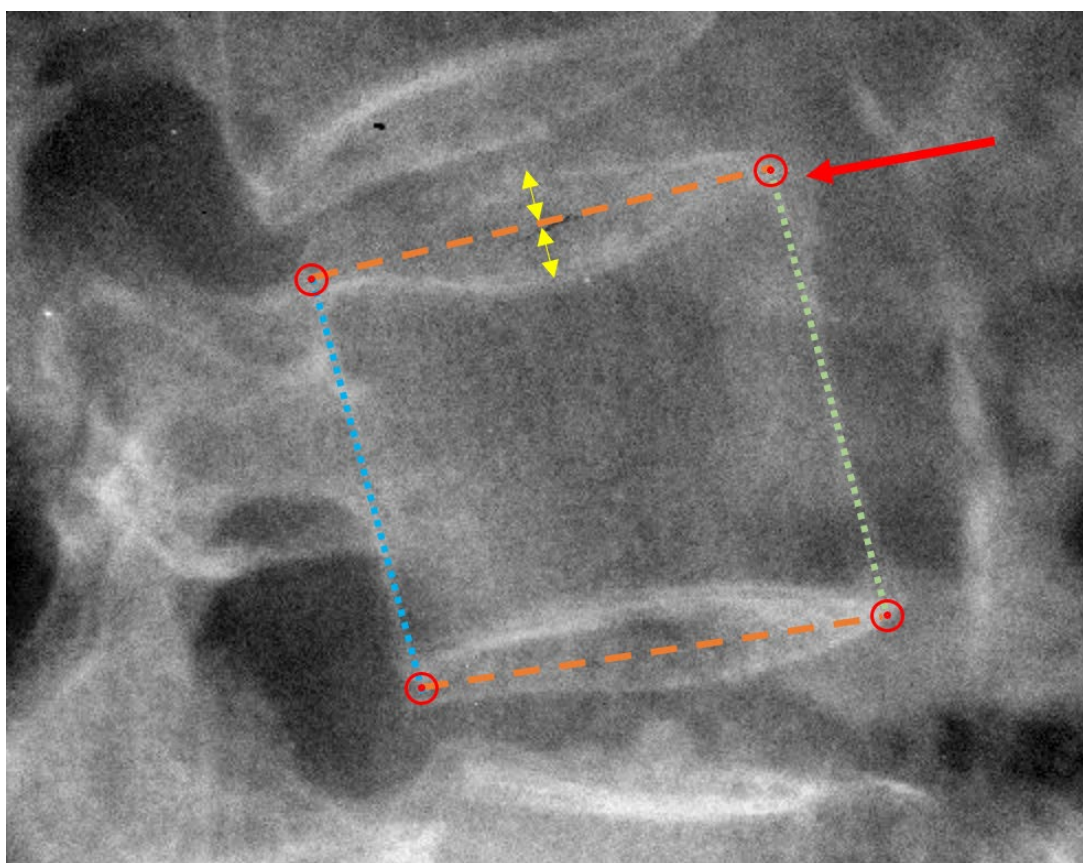
Details of vertebral landmark placement

The standardized landmark placement was focused on the mid-sagittal plane of each vertebral body (Figure 1), excluding any residual uncinate processes, and placing landmarks that identify the corners of the vertebral bodies prior to any osteophyte formation⁽¹⁻³⁾. The posterior superior landmarks are placed so as to represent the endplate as it would appear on a mid-sagittal slice of a CT exam, rather than up at the top of any posterior ridges on the upper endplate, similar to logic described in Keynan et al and Quint et al.^(4,5).

When the X-ray beam path through a vertebra is not perpendicular to the mid-sagittal plane of the vertebra, the left and right rims of the vertebral endplates (as seen in Figure 1), and the left and right aspects of the posterior vertebral body may be seen on the X-ray.^(3,6) The mid-sagittal plane is assumed to be midway between the radiographic shadows of the left and right endplate rims, and midway between the radiographic shadows of the left and right aspects of the posterior wall. A similar approach was described by Quint et al.⁽⁵⁾ Multiple examples of landmark placement are provided online¹ since disc space metrics calculated from landmarks are dependent on the nuances of how landmarks are placed, and it is therefore important to appreciate landmark placement details.

¹ <https://www.dropbox.com/sh/qzrocrh86goxarx/AAADRxs5HoEjVdRcGIwA4bela?dl=0>

Figure 1: Details of anatomic landmark placement. Dashed lines show the estimated mid-sagittal plane of the superior and inferior endplates, identified as bisecting the radiographic shadows of the left and right sides of the endplates (yellow arrows). The red circles show the four landmarks used to measure vertebral body morphology and disc space. The red arrow points to an anterior osteophyte that is ignored when placing landmarks. The dotted green and blue lines show the anterior and posterior aspects of the vertebral body.



Disc space and sagittal plane offset (SPO) metrics from the NHANES-II lumbar spine radiographs:

Table 1: Descriptive statistics for Anterior disc heights (ADH), before and after [in brackets] trimming.

Level	N	Mean	Std Dev	CV	Median	10th	95th	Skew	Kurtosis
L1L2	7020 [3487]	25.47 [25.84]	4.46 [2.75]	0.17 [0.11]	25.53 [25.82]	20.04 [22.14]	32.48 [30.44]	-0.28 [0.06]	3.87 [2.31]
L2L3	7284 [3675]	28.75 [29.36]	4.89 [2.94]	0.17 [0.10]	29.07 [29.43]	22.68 [25.29]	36.16 [34.19]	-0.56 [-0.05]	4.25 [2.32]
L3L4	7297 [3842]	32.03 [32.71]	5.31 [3.11]	0.17 [0.09]	32.36 [32.62]	25.47 [28.59]	39.90 [37.91]	-0.67 [-0.01]	4.47 [2.42]
L4L5	7163 [3874]	36.76 [7.79]	7.01 [3.99]	0.19 [0.11]	37.65 [37.95]	27.54 [32.34]	46.37 [44.07]	-0.76 [-0.22]	4.07 [2.50]
L5S1	6682 [3573]	41.38 [42.32]	9.30 [5.84]	0.22 [0.14]	42.20 [42.57]	28.41 [34.29]	55.44 [51.56]	-0.36 [-0.18]	2.96 [2.39]

Table 2: Descriptive statistics for Posterior disc heights (PDH), before and after [in brackets] trimming.

Level	N	Mean	Std Dev	CV	Median	10th	95th	Skew	Kurtosis
L1L2	7020 [3487]	17.71 [17.97]	3.17 [2.17]	0.18 [0.12]	17.67 [17.88]	13.77 [15.03]	22.97 [21.80]	0.05 [0.15]	3.57 [2.25]
L2L3	7284 [3675]	17.12 [17.55]	3.66 [2.44]	0.21 [0.14]	17.13 [17.47]	12.46 [14.38]	23.19 [21.72]	-0.03 [0.10]	3.05 [2.27]
L3L4	7297 [3842]	17.51 [18.01]	3.91 [2.59]	0.22 [0.14]	17.58 [18.03]	12.43 [14.48]	23.73 [22.31]	-0.10 [-0.01]	3.24 [2.23]
L4L5	7163 [3874]	17.36 [18.01]	4.55 [3.06]	0.26 [0.17]	17.57 [18.14]	11.28 [13.75]	24.56 [22.95]	-0.15 [-0.08]	3.05 [2.26]
L5S1	6682 [3573]	13.38 [13.83]	4.73 [3.29]	0.35 [0.24]	13.22 [13.83]	7.56 [9.39]	21.33 [19.50]	0.17 [0.13]	3.30 [2.21]

Table 3: Descriptive statistics for Average Disc Height (AvgDH), before and after [in brackets] trimming.

Level	N	Mean	Std Dev	CV	Median	10th	90th	Skew	Kurtosis
L1L2	7020 [3436]	21.59 [21.94]	3.16 [2.08]	0.15 [0.09]	21.61 [21.83]	17.70 [19.23]	26.77 [25.45]	-0.23 [0.15]	3.99 [2.35]
L2L3	7284 [3638]	22.94 [23.49]	3.63 [2.29]	0.16 [0.10]	23.18 [23.48]	18.37 [20.37]	28.55 [27.29]	-0.51 [0.04]	3.95 [2.36]
L3L4	7297 [3811]	24.77 [25.39]	3.96 [2.47]	0.16 [0.10]	25.05 [25.36]	19.75 [22.13]	30.59 [29.41]	-0.65 [0.00]	4.25 [2.35]
L4L5	7163 [3848]	27.06 [27.94]	5.00 [3.03]	0.18 [0.11]	27.71 [28.03]	20.28 [23.86]	33.97 [32.69]	-0.77 [-0.18]	3.81 [2.46]
L5S1	6682 [3505]	27.38 [28.16]	5.74 [3.77]	0.21 [0.13]	28.01 [28.35]	19.24 [22.93]	35.96 [34.11]	-0.35 [-0.17]	2.96 [2.49]

Table 4: Descriptive statistics for Disc Angle (DA), before and after [in brackets] trimming.

Level	N	Mean	Std Dev	CV	Median	10th	90th	Skew	Kurtosis
L1L2	7020 [3487]	4.59 [4.64]	2.63 [1.53]	0.57 [0.33]	4.61 [4.66]	1.27 [2.57]	7.90 [6.70]	-0.11 [-0.03]	3.40 [2.25]
L2L3	7284 [3675]	6.87 [6.96]	2.78 [1.64]	0.40 [0.24]	6.87 [6.96]	3.38 [4.70]	10.35 [9.20]	-0.14 [-0.04]	3.25 [2.20]
L3L4	7297 [3842]	8.45 [8.54]	2.88 [1.64]	0.34 [0.19]	8.52 [8.55]	4.91 [6.33]	12.04 [10.75]	-0.17 [-0.04]	3.61 [2.26]
L4L5	7163 [3874]	11.15 [11.36]	3.66 [2.07]	0.33 [0.18]	11.37 [11.46]	6.42 [8.52]	15.51 [14.06]	-0.29 [-0.20]	3.56 [2.31]
L5S1	6682 [3573]	16.19 [16.49]	5.37 [3.20]	0.33 [0.19]	16.41 [16.60]	9.11 [12.08]	23.04 [20.70]	-0.19 [-0.12]	2.95 [2.19]

Table 5: Descriptive statistics for Anterior Sagittal Plane Offset (ASPO), before and after [in brackets] trimming.

Level	N	Mean	StdDev	CV	Median	5th	95th	Skew	Kurtosis
L1L2	7020 [3487]	-5.97 [-5.72]	2.48 [1.64]	-0.42 [-0.29]	-5.81 [-5.65]	-10.31 [-8.63]	-2.29 [-3.10]	-0.44 [-0.27]	4.44 [2.64]
L2L3	7284 [3675]	-7.40 [-7.22]	2.86 [1.86]	-0.39 [-0.26]	-7.35 [-7.20]	-12.23 [-10.36]	-3.02 [-4.14]	0.00 [-0.08]	4.93 [2.50]
L3L4	7297 [3842]	-6.38 [-6.46]	3.50 [2.10]	-0.55 [-0.33]	-6.47 [-6.47]	-11.65 [-9.95]	-0.75 [-2.85]	0.70 [0.05]	6.20 [2.66]
L4L5	7163 [3874]	-5.55 [-6.05]	5.33 [3.05]	-0.96 [-0.50]	-6.32 [-6.34]	-12.37 [-10.55]	4.01 [-0.16]	1.52 [0.58]	8.00 [3.18]
L5S1	6682 [3573]	-9.71 [-10.19]	6.33 [3.70]	-0.65 [-0.36]	-10.31 [-10.43]	-18.50 [-15.90]	0.89 [-3.40]	1.31 [0.33]	7.60 [2.80]

Table 6: Descriptive statistics for Posterior Sagittal Plane Offset (PSPO), before and after [in brackets] trimming.

Level	N	Mean	StdDev	CV	Median	5th	95th	Skew	Kurtosis
L1L2	7020 [3487]	2.64 [2.55]	2.32 [1.53]	0.88 [0.60]	2.56 [2.52]	-0.84 [0.05]	6.46 [5.21]	0.01 [0.15]	5.60 [2.54]
L2L3	7284 [3675]	3.97 [4.02]	2.48 [1.66]	0.62 [0.41]	3.95 [4.01]	0.08 [1.27]	7.90 [6.88]	-0.29 [0.04]	5.61 [2.45]
L3L4	7297 [3842]	4.11 [4.34]	3.06 [1.89]	0.74 [0.44]	4.31 [4.39]	-0.79 [1.13]	8.49 [7.39]	-1.08 [-0.15]	7.74 [2.50]
L4L5	7163 [3874]	3.98 [4.53]	4.75 [2.80]	1.19 [0.62]	4.74 [4.89]	-4.65 [-0.83]	9.94 [8.56]	-1.86 [-0.59]	10.38 [3.05]
L5S1	6682 [3573]	5.99 [6.40]	5.67 [3.28]	0.95 [0.51]	6.64 [6.56]	-3.24 [0.59]	13.24 [11.52]	-1.93 [-0.34]	11.27 [2.77]

Table 7: Descriptive statistics for Ventral disc heights (VDH), before and after [in brackets] trimming.

Level	N	Mean	StdDev	CV	Median	10th	95th	Skew	Kurtosis
L1L2	7020 [3487]	26.52 [26.86]	4.68 [2.87]	0.18 [0.11]	26.62 [26.84]	20.88 [22.98]	34.01 [31.62]	-0.26 [0.05]	3.93 [2.32]
L2L3	7284 [3675]	29.93 [30.48]	5.12 [3.07]	0.17 [0.10]	30.16 [30.53]	23.61 [26.29]	37.78 [35.56]	-0.50 [-0.03]	4.22 [2.33]
L3L4	7297 [3842]	32.61 [33.26]	5.46 [3.17]	0.17 [0.10]	32.95 [33.19]	25.92 [29.07]	40.77 [38.53]	-0.63 [-0.02]	4.43 [2.43]
L4L5	7163 [3874]	36.88 [37.93]	7.18 [4.06]	0.19 [0.11]	37.76 [38.11]	27.45 [32.37]	47.00 [44.25]	-0.70 [-0.22]	3.98 [2.54]
L5S1	6682 [3573]	41.62 [42.69]	9.50 [5.95]	0.23 [0.14]	42.43 [42.95]	28.62 [34.55]	55.95 [51.95]	-0.34 [-0.18]	2.95 [2.38]

Table 8: Descriptive statistics for Dorsal disc heights (DDH), before and after [in brackets] trimming.

Level	N	Mean	StdDev	CV	Median	10th	95th	Skew	Kurtosis
L1L2	7020 [3487]	18.40 [18.63]	3.25 [2.23]	0.18 [0.12]	18.35 [18.56]	14.38 [15.64]	23.80 [22.52]	0.06 [0.14]	3.63 [2.26]
L2L3	7284 [3675]	17.81 [18.20]	3.69 [2.46]	0.21 [0.14]	17.79 [18.13]	13.12 [14.99]	23.92 [22.38]	-0.03 [0.09]	3.06 [2.27]
L3L4	7297 [3842]	17.87 [18.36]	3.87 [2.56]	0.22 [0.14]	17.94 [18.38]	12.87 [14.84]	23.98 [22.62]	-0.12 [-0.02]	3.29 [2.23]
L4L5	7163 [3874]	17.53 [18.20]	4.53 [3.04]	0.26 [0.17]	17.78 [18.29]	11.50 [13.99]	24.62 [23.14]	-0.19 [-0.09]	3.10 [2.27]
L5S1	6682 [3573]	13.85 [14.34]	4.53 [3.21]	0.33 [0.22]	13.74 [14.31]	8.17 [10.02]	21.48 [19.87]	0.13 [0.10]	3.19 [2.21]

Table 9: Descriptive statistics for Mid-Plane Disc Angles (MPA), before and after [in brackets] trimming.

Level	N	Mean	StdDev	CV	Median	10th	90th	Skew	Kurtosis
L1L2	7020 [3487]	2.65 [2.74]	3.27 [1.94]	1.23 [0.71]	2.70 [2.81]	-1.35 [0.09]	6.75 [5.38]	-0.17 [-0.03]	3.67 [2.20]
L2L3	7284 [3675]	6.31 [6.34]	3.41 [2.08]	0.54 [0.33]	6.32 [6.32]	2.06 [3.50]	10.56 [9.17]	-0.11 [-0.02]	3.37 [2.13]
L3L4	7297 [3842]	9.16 [9.13]	3.59 [2.16]	0.39 [0.24]	9.15 [9.12]	4.67 [6.15]	13.62 [12.07]	0.01 [0.03]	3.43 [2.18]
L4L5	7163 [3874]	15.16 [15.17]	4.46 [2.67]	0.29 [0.18]	15.32 [15.25]	9.55 [11.46]	20.69 [18.67]	-0.11 [-0.06]	3.51 [2.17]
L5S1	6682 [3573]	24.07 [24.42]	5.71 [3.49]	0.24 [0.14]	24.44 [24.65]	16.44 [19.52]	31.04 [29.01]	-0.27 [-0.19]	3.11 [2.21]

Table 10: Descriptive statistics for Centroid Sagittal Plane Offset (CSPO), before and after [in brackets] trimming.

Level	N	Mean	StdDev	CV	Median	5th	95th	Skew	Kurtosis
L1L2	7020 [3487]	-5.18 [-4.93]	2.62 [1.80]	-0.51 [-0.37]	-5.01 [-4.85]	-9.66 [-8.10]	-1.35 [-2.06]	-0.52 [-0.23]	4.34 [2.44]
L2L3	7284 [3675]	-7.43 [-7.24]	2.73 [1.86]	-0.37 [-0.26]	-7.29 [-7.14]	-11.98 [-10.46]	-3.26 [-4.20]	-0.19 [-0.14]	5.03 [2.48]
L3L4	7297 [3842]	-8.00 [-8.06]	3.10 [1.96]	-0.39 [-0.24]	-8.11 [-8.05]	-12.67 [-11.24]	-3.15 [-4.76]	0.76 [0.06]	6.71 [2.46]
L4L5	7163 [3874]	-8.20 [-8.66]	4.74 [2.85]	-0.58 [-0.33]	-8.94 [-9.01]	-13.94 [-12.81]	0.44 [-3.17]	1.84 [0.61]	9.88 [3.00]
L5S1	6682 [3573]	4.21 [4.12]	6.18 [4.07]	1.47 [0.99]	3.77 [3.95]	-5.12 [-2.56]	14.20 [11.23]	0.93 [0.12]	6.48 [2.55]

Proportion of levels with abnormalities. This includes all lumbar spines from the NHANES-II study. Abnormalities were identified by standardized (Z) scores in units of std dev from average normal.

Table 11: Proportion with disc heights > normal

Level	zADH>2	zADH>3	zPDH>2	zPDH>3	zVDH>2	zVDH>3	zDDH>2	zDDH>3
L1L2	8.4	2.4	7.2	1.7	8.7	2.5	7.6	1.8
L2L3	7.4	1.7	7.5	1.7	7.9	2.1	7.6	1.8
L3L4	7.4	1.7	6.6	1.7	8.2	2.0	6.4	1.7
L4L5	6.5	1.4	6.1	1.2	7.0	1.7	5.9	1.1
L5S1	7.0	1.2	7.3	1.8	7.0	1.2	6.5	1.5

Table 12: Proportions with disc Heights < normal

Level	zADH<-2	zADH<-3	zPDH<-2	zPDH<-3	zVDH<-2	zVDH<-3	zDDH<-2	zDDH<-3
L1L2	11.2	4.1	9.1	2.3	11.0	4.0	8.8	2.3
L2L3	12.4	5.2	11.1	3.2	12.0	5.0	10.6	3.2
L3L4	13.5	5.9	11.6	3.6	13.1	5.9	11.5	3.6
L4L5	14.4	7.6	12.0	3.6	14.5	7.6	12.4	3.8
L5S1	13.9	5.0	8.9	1.8	14.1	5.1	9.0	1.8

Table 13: Proportions with disc angles > normal

Level	zDA>2	zDA>3	zMPA>2	zMPA>3
L1L2	11.4	3.8	10.7	3.1
L2L3	10.8	3.2	10.2	3.2
L3L4	11.3	3.8	11.0	3.9
L4L5	10.0	2.7	10.6	3.1
L5S1	10.4	2.7	8.8	2.0

Table 14: Proportions with disc angles < normal

Level	zDA<-2	zDA<-3	zMPA<-2	zMPA<-3
L1L2	12.2	4.3	11.2	3.9
L2L3	11.9	4.3	10.8	3.5
L3L4	12.2	4.8	10.5	3.3
L4L5	13.9	5.9	11.2	4.1
L5S1	13.4	4.8	12.8	4.8

Table 15: Proportions with SPO > normal

Level	zASPO>2	zASPO>3	zPSPO>2	zPSPO>3	zCSPO>2	zCSPO>3
L1L2	5.6	1.1	9.3	3.0	4.9	0.8
L2L3	7.0	1.9	7.7	2.1	6.1	1.4
L3L4	8.9	4.0	6.3	1.7	7.9	3.5
L4L5	11.3	5.8	4.3	0.6	9.9	5.5
L5S1	10.9	5.0	5.7	1.0	7.7	3.6

Table 16: Proportions with SPO < normal

Level	zASPO<-2	zASPO<-3	zPSPO<-2	zPSPO<-3	zCSPO<-2	zCSPO<-3
L1L2	10.6	4.0	6.4	1.6	9.6	3.4
L2L3	9.8	3.5	7.5	2.3	9.2	3.0
L3L4	8.7	2.9	9.1	4.0	7.4	2.0
L4L5	5.6	1.0	10.7	5.6	3.8	0.6
L5S1	7.2	1.2	9.3	4.7	6.8	1.6

References

1. Hurxthal LM. Measurement of anterior vertebral compressions and biconcave vertebrae. *Am J Roentgenol Radium Ther Nucl Med*. Jul 1968;103(3):635-44.
2. Diacinti D, Guglielmi G. Vertebral morphometry. *Radiologic clinics of North America*. May 2010;48(3):561-75.
3. Frobin W, Brinckmann P, Biggemann M, Tillotson M, Burton K. Precision measurement of disc height, vertebral height and sagittal plane displacement from lateral radiographic views of the lumbar spine. *Clin Biomech (Bristol, Avon)*. 1997;12 Suppl 1:S1-S63.
4. Keynan O, Fisher CG, Vaccaro A, Fehlings MG, Oner FC, Dietz J, et al. Radiographic measurement parameters in thoracolumbar fractures: a systematic review and consensus statement of the spine trauma study group. *Spine (Phila Pa 1976)*. 10.1097/01.brs.0000201261.94907.0d doi ;00007632-200603010-00025 pii Mar 1 2006;31(5):E156-65.
5. Quint DJ, Tuite GF, Stern JD, Doran SE, Papadopoulos SM, McGillicuddy JE, et al. Computer-assisted measurement of lumbar spine radiographs. *Academic radiology*. 1997;4(11):742-52.
6. Brinckmann P, Frobin W, Biggemann M, Hilweg D, Seidel S, Burton K, et al. Quantification of overload injuries to thoracolumbar vertebrae and discs in persons exposed to heavy physical exertions or vibration at the work-place: The shape of vertebrae and intervertebral discs—study of a young, healthy population and a middle-aged control group. *Clinical Biomechanics*. 1994;9:S3-S83.

Supplemental Appendix 2: Measurement of spondylolisthesis in flexion and extension

Background and Methods

Although disc heights and disc angles significantly explained some of the variance in Sagittal Plane Offset (SPO) metrics in the NHANES-II data, the influence was relatively small (best R^2 for anterior SPO was 0.56) and much of the variability in SPO was left unexplained. The NHANES-II radiographs were obtained with subjects side-lying with relatively little variation in L1-S1 angle expected between subjects, compared to what can be expected in flexion-extension radiographs obtained with good subject effort. Spondylolisthesis may sometimes be assessed from flexion and extension (FE) radiographs in clinical practice and research studies¹. Therefore, Anterior SPO (ASPO), Posterior SPO (PSPO) and Centroid SPO (CSPO) was also analyzed in flexion-extension radiographs of asymptomatic volunteers.

Some (162) of the FE radiographs were described and analyzed in a prior publication.² An additional 211 FE radiographs were also analyzed. The first 162 FE studies had been obtained using a seated protocol², while the second 211 had been obtained with an upright protocol using a standard walker for support. Each of the 211 subjects watched a short training video before the X-rays were collected¹. All 373 subjects consented to the IRB approved studies, were asymptomatic, and had never had medical treatment for a back disorder.

The FE studies were analyzed using the same series of neural networks and coded logic (Spine CAMP™, Medical Metrics, Inc., Houston, TX) as with the NHANES-II X-rays. In addition, three highly experienced musculoskeletal radiologists assessed disc degeneration at every level using the Kellgren-Lawrence (KL) grading system.^{3,4} The first two radiologists assessed all radiographs, and the third

¹ <https://youtu.be/KaEQP57qWgU> English language version

https://youtu.be/Day_wvEG-yI Spanish language version

adjudicated if the first two did not agree on the KL grade. The nine SPO and disc metrics were calculated from the landmarks using only those levels with no more than KL grade 1 (doubtful) disc degeneration. Regression analysis was used to obtain an equation to predict ASPO from level, ADH, and PDH. (Stata ver 15, College Station, TX).

Results

Applying the normative data from the NHANES-II study (where all subjects were imaged in a side-lying position) to the 373 flexion-extension radiographs, 45% of the flexion and extension radiographs have standardized ASPO metrics >2 or <-2 Std Dev from average normal ASPO. Since 45% of asymptomatic levels can't have abnormal SPO, this documents that normative SPO reference data established using the NHANES-II data should NOT be used with flexion and extension radiographs, at least with a threshold of >2 or <-2 Std Dev from average normal ASPO to classify SPO as abnormal.

To establish a method for measuring SPO in the presence of flexion or extension, the disc height, disc angle, and SPO data for the flexion and extension radiographs were trimmed to avoid including abnormalities. Trimming of the flexion-extension data was based on both SPO metrics ($<5^{\text{th}}$ or $>95^{\text{th}}$ percentiles excluded) and the radiologist Kellgren-Lawrence (KL) grades to identify abnormal or degenerated levels. All levels in the flexion-extension data were trimmed where ASPO, PSPO, or CSPO was abnormal or where radiographic disc degeneration was $>$ “doubtful” (KL grade 1). Based on the radiologist's grading of disc degeneration combined with the SPO metrics, 25% of levels had degeneration but normal SPO, 11.3% had abnormal SPO but no degeneration, and 7.8% had both degeneration and abnormal SPO. Thus, 44.2% of levels were excluded from establishing an equation to predict normal SPO from disc heights and disc angles.

Analysis of variance revealed that the combination of ADH, PDH, and DA predicted 72% of the variation in PSPO and 85% of the variation in ASPO. The combination of VDH, DDH and MPA

predicted 71% of the variation in CSPO. Since the R^2 for ASPO was the highest, a regression equation was developed for predicting normal ASPO from ADH, PDH, and DA. Using the trimmed data, the Equation 1:

$$\text{Predicted ASPO} = C1 + \text{ADH} * C2 + \text{PDH} * C3 + \text{DA} * C4 + C5$$

Where:

$$C1 = 0 \text{ if L1-L2; } 0.173 \text{ if L2L3; } 1.41 \text{ if L3L4; } 2.54 \text{ if L4L5, } 0.647 \text{ if L5S1}$$

$$C2 = 1.53$$

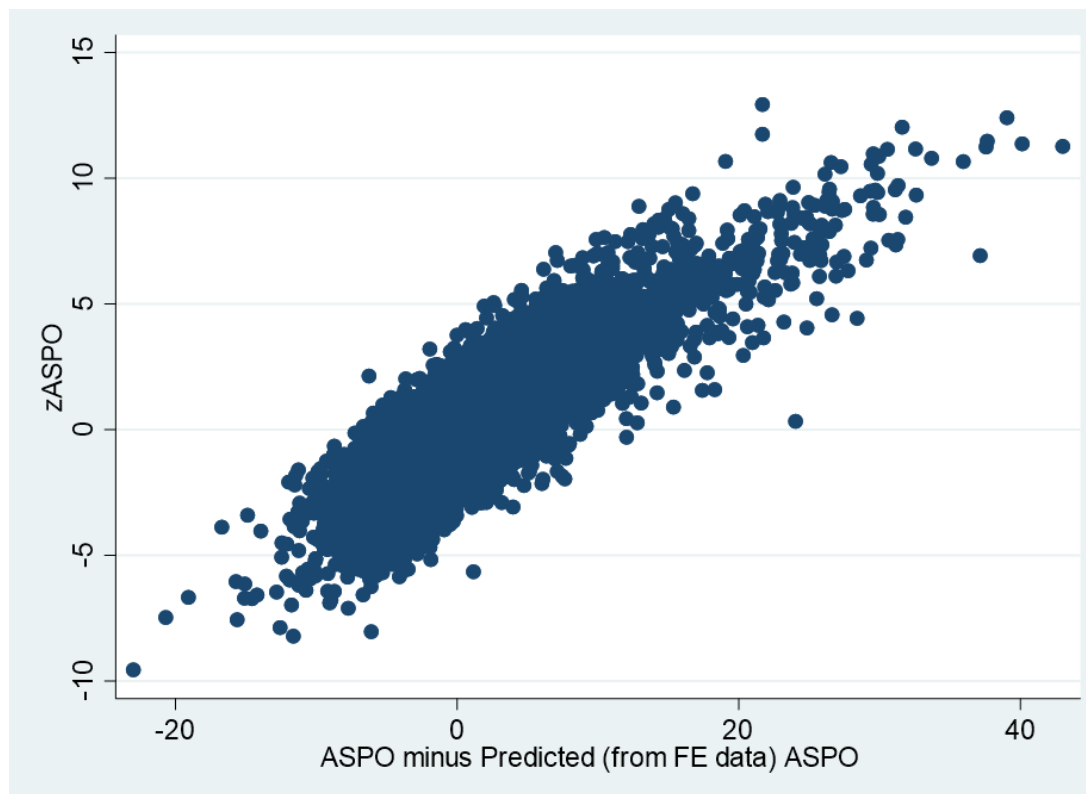
$$C3 = -1.50$$

$$C4 = -3.19$$

$$C5 = -3.12$$

As a check of the regression equation, the predicted ASPO was linearly ($P < 0.0001$) correlated to the actual ASPO for the trimmed data ($R^2 = 0.85$). The average difference between the predicted ASPO and the actual ASPO, using only the trimmed NHANES-II data, was 0.014 % endplate width (std dev 2.1). The differences ranged from -6.3 to 11.4 % endplate width. Including all of the NHANES-II data, the difference between the predicted (from Equation 1) and actual ASPO was strongly ($R^2 = 0.64$) and linearly ($P < 0.0001$) correlated to the standardized ASPO metric (Figure 1). A stronger correlation would not be expected, since the predicted ASPO is the predicted **normal** ASPO, and actual ASPO in the NHANES-II study is not always normal. It is the difference between the actual and predicted (normal) ASPO that could be of clinical value. This supports the potential value of the ASPO prediction equation from the asymptomatic flexion-extension X-Rays. Further research will be required to determine which metric is best for classifying SPO as normal versus abnormal: the standardized ASPO, PSPO, or CSPO metric from the NHANES-II study, or the prediction equation derived from the flexion-extension studies. Nevertheless, this supports the hypothesis that it is necessary to account for disc heights and disc angle when predicting what normal SPO should be, particularly if there is a lot of flexion or extension.

Figure 1: On the X-axis is the difference (units of percent endplate width) between the actual ASPO and the predicted ASPO (using the prediction equation derived from flexion-extension radiographs). On the Y-axis is the standardized ASPO (standard deviations from average ASPO in the trimmed NHANES-II data).



ASPO predicted using the flexion-extension X-rays can also be expressed as a standardized metric (difference from average normal). Since the predicted ASPO is calculated from an equation, a different approach to standardization is used. To obtain the denominator needed to calculate a standardized predicted ASPO, the standard error of the forecast was calculated after the regression. The standard error of the forecast is the standard error of the point prediction for 1 observation. There was minimal difference between levels, so one average value for all levels was used (1.87). A standardized SPO that accounts for disc heights and angles was then calculated as the actual ASPO minus the predicted

ASPO divided by the standard error of the forecast. This is referred to as the spondylolisthesis index (SI) to distinguish it from the other SPO metrics described in the NHANES-II study.

To assess the difference in SPO between flexion and extension, the SI was calculated for every level for both the flexion and extension x-rays. Table 1 provides the median, 5th and 95th percentile differences between standardized (using NHANES-II data) ASPO, PSPO, and CSPO calculated in flexion versus extension. The difference in SI in flexion versus extension is also provided. Very large differences were observed between flexion and extension in the NHANES-II standardized ASPO, PSPO, and CSPO metrics. This also supports that the SPO metrics from the NHANES-II study should NOT be used for flexion-extension radiographs. Smaller differences were observed in the SI. Note that there were some cases in the flexion-extension X-rays of asymptomatic volunteers with clear spondylolisthesis. The data document that the median SI changes by approximately 1 between flexion and extension in an asymptomatic population, and that there are some levels with minimal change and others with substantial change. Further research is needed to determine if this can be of clinical value.

Table 1: Differences in the standardized SPO metrics when the metric in flexion is compared to the same metric in extension. Note the large variations in ASPO, PSPO, and CSPO compared to the SI.

Level	Diff ASPO Flex v Ext			Diff PSPO Flex v Ext			Diff CSPO Flex v Ext			Diff SI Flex v Ext		
	50th	5th	95th	50th	5th	95th	50th	5th	95th	50th	5th	95th
L1L2	3.2	1.2	5.7	3.1	0.8	5.0	1.8	0.3	3.3	0.5	0.1	1.9
L2L3	4.2	2.1	6.4	3.6	1.5	5.8	2.4	1.0	4.2	0.6	0.0	2.0
L3L4	4.3	2.2	6.6	3.6	1.5	5.8	2.8	1.3	4.3	1.0	0.1	2.8
L4L5	3.5	1.2	5.6	2.6	0.6	4.2	2.0	0.6	3.2	1.3	0.1	3.6
L5S1	1.8	0.2	4.5	0.6	0.1	2.2	0.4	0.0	1.6	0.8	0.1	2.8

References

1. Mummaneni PV, Bisson EF, Kerezoudis P, et al. Minimally invasive versus open fusion for Grade I degenerative lumbar spondylolisthesis: analysis of the Quality Outcomes Database. *Neurosurgical focus*. 2017;43(2):E11.
2. Staub BN, Holman PJ, Reitman CA, Hipp JA. Sagittal plane lumbar intervertebral motion during seated flexion-extension in 658 asymptomatic, non-degenerated levels. *J Neurosurgery Spine*. 2015;23(6):731-738. doi:10.3171/2015.3.SPINE14898
3. Kellgren J, Lawrence J. Radiological assessment of osteo-arthritis. *Annals of the Rheumatic Diseases*. 1957;16(4):494.
4. Kellgren J, Lawrence J. Atlas of standard radiographs: the epidemiology of chronic rheumatism. Vol. 2. Oxford: Blackwell; 1963.

Supplemental Appendix 3: Analysis of errors due to out-of-plane imaging

Materials and Methods

To document expected variability in the metrics due to variations in radiographic projections that could occur in clinical practice, digitally reconstructed radiographs (DRR) were used. Thin-slice (1 mm or less), anonymized computed tomography exams of the lumbar spine of 26 individuals were interpolated to 0.2 mm isotropic resolution. The interpolation was done to achieve higher resolution DRRs. The three-dimensional (3D) coordinates of the four anatomic landmarks described in Supplementary Appendix 1, Figure 1 were manually digitized in the mid-sagittal plane of each vertebra from L1 to S1. Landmark position was identified using intersecting axial, coronal, and sagittal slices. Landmarks were thus placed in the mid-sagittal plane at clear and well-defined locations. The image processing and landmark placement were completed using Slicer 3D.¹ The nine disc and SPO metrics described in the paper were calculated directly from the 3D landmark coordinates and used as the “gold standard” to assess accuracy of measurements from the DRRs.

Custom python code was developed to create 2D, lateral, digitally reconstructed radiographs from the 3D CT data using the Plastimatch DRR function². A 40 inch source-to-image distance was used for all X-rays. The base x-ray was centered on the centroid of the L3 vertebra (this is the baseline isocenter). The distance between the center of L3 and the image was set to $\frac{1}{2}$ the estimated width of the hips. The projection matrices used by Plastimatch DRR to create the simulated X-rays were then used to calculate, from the known 3D coordinates, the precise coordinates of each landmark on each 2D simulated X-ray. In addition to the baseline DRR, forty additional DRRs were generated with the following variations:

- random beam tilts of between ± 10 deg in each of the sagittal and axial planes
- random displacements from the baseline isocenter of ± 2 endplate widths in the AP and LR directions, plus ± 3 endplate width shifts in the cranial-caudal direction

- random image rotations between ± 24 deg

These variations resulted in a wide range of radiographic projections, and represent a digital version of the methodology used in earlier studies that addressed the effect of radiographic projection on disc height measurements.³⁻⁵ The landmarks in every image were precisely calculated so all 2D landmark coordinates corresponded to the true 3D coordinates. This allows for documenting the effect of errors in vertebral morphology metrics that are due to radiographic projection, without uncertainty in landmark placement. These precisely calculated landmarks were analyzed to obtain the nine disc height, disc angle, and SPO metrics described in the paper. There were thus 41 measurements for every disc level (L1-L2 to L5-S1) for each of the 26 independent CT exams.

Although the coordinates of landmarks in every simulated X-ray were precisely calculated based on geometry, the quality of the simulated radiographs was below clinical standards, did not include the field-of-view normally found in a lateral spine X-ray, and were therefore not appropriate for use toward validating the accuracy of neural networks in obtaining landmarks. It should also be appreciated that the simulated radiographs with no beam tilt or offset cannot be assumed to be the optimal radiograph, since not all CT exams were obtained with perfectly oriented spines and some spines had frontal plane curvatures in the lumbar spine.

Results

Table 1 provides the median, 5th and 95th percentile errors in the metrics. These are errors solely due to variability in radiographic projection with no error in landmark digitization. Standardized metrics (units of Std Dev from average normal) are reported in Table 1 to facilitate interpretation of the data and allow for pooling of data for all levels. For example, the median error in ASPO, due to random variations in radiographic projection, was 0.19 std dev from average ASPO in healthy discs. The maximum error in ASPO that occurred due to variations in radiographic projection was 1.1 std dev.

Discussion

In general, errors due to variability in radiographic projection are small. However, large errors can occur with severe out-of-plane imaging. Disc angles were the least effected, followed by disc heights and then SPO. The worst errors were found in CT exams of scoliotic spines (eg Figure 1). With the spine in Figure 1, optimal imaging of the L2L3 level would require substantial beam tilt, and that amount of beam tilt would yield poor quality imaging of the L4L5 level.

Conclusion

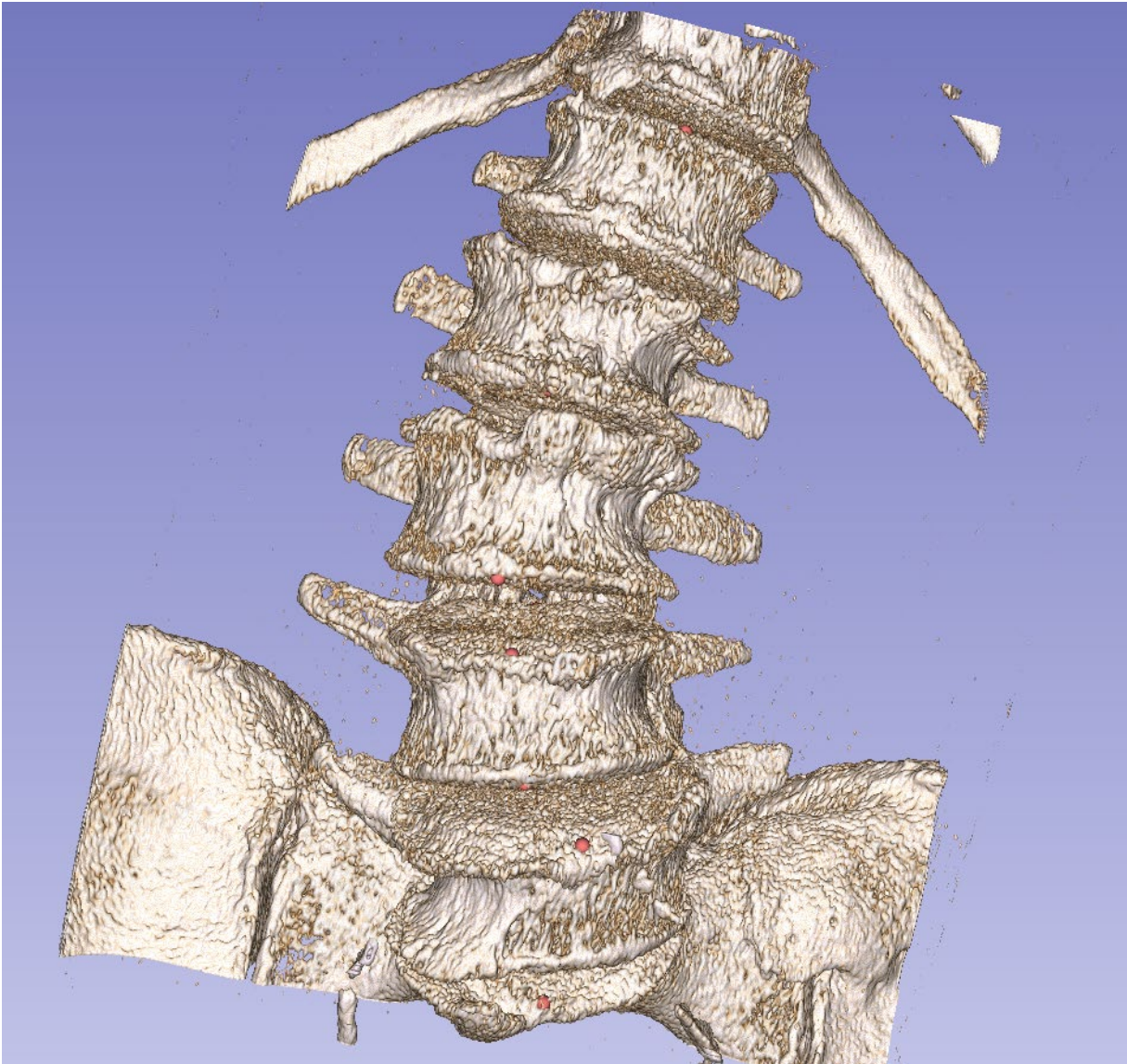
This experiment documents the need for caution when interpreting sagittal plane disc and SPO metrics when the radiographic projection is substantially out-of-plane. The images can be out-of-plane both with respect to the frontal and axial planes. A method to determine when there is too much out-of-plane is challenging. This requires both a definition of how much error is too much, and a method to detect out-of-plane that would result in too much error. Neural networks can likely be trained to recognize excessive out-of-plane once error tolerances are defined.

Table 1 : Descriptive statistics to help assess the effect of radiographic projection on error in the metrics.

The absolute error for each metric was measured by comparing metrics from 2D simulated radiographs to the known measurement made from 3D data. Errors were analyzed using standardized (“z”) scores in units of std devs from average in healthy discs. Over 5,000 measurements were available for each metric (26 CT exams * 5 levels * 41 variations in radiographic projection).

Variable	Median	5th %	95th %
zADH	0.13	0.012	0.92
zPDH	0.12	0.0089	0.79
zDA	0.093	0.0061	0.6
zASPO	0.19	0.013	1.12
zPSPO	0.14	0.009	0.90
zVDH	0.13	0.010	0.86
zDDH	0.12	0.0094	0.76
zMPA	0.12	0.0099	0.47
zCSPO	0.43	0.026	2.13

Figure 1: Three-dimensional reconstruction of one of the spines in the simulated X-ray study. Considering how a lateral radiograph of the lumbar spine would appear, helps to understand how disc heights and angles could change as the x-ray beam angle changes in the frontal plane.



References

1. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging*. 2012;30(9):1323-1341.
doi:<https://doi.org/10.1016/j.mri.2012.05.001>
2. Shackelford JA, Shusharina N, Verberg J, et al. Plastimatch 1.6-current capabilities and future directions. 2012:
3. Andersson G, Schultz A, Nathan A, Irstam L. Roentgenographic measurement of lumbar intervertebral disc height. *Spine*. 1981;6(2):154-158.
4. Saraste H, Brostrom LA, Aparisi T, Axdorph G. Radiographic measurement of the lumbar spine. A clinical and experimental study in man. *Spine*. 1985;10(3):236-241. NOT IN FILE.
5. Frobin W, Brinckmann P, Biggemann M, Tillotson M, Burton K. Precision measurement of disc height, vertebral height and sagittal plane displacement from lateral radiographic views of the lumbar spine. *Clinical Biomechanics*. 1997;12:S1-S63.