CLINICAL ARTICLE

A Preoperative Predictive Model of Lower Lumbar Spine Instability Based on Three-Dimensional Computed Tomography: A Retrospective Case–Control Pilot Study

Yanbing Kao, MD[1], Zijing Liu, MBBS[2], Jiali Leng, MBBS[3], Zhigang Qu, MD[1], QingXu Song, MD[1], Yi Liu, MD[1], Zhenyu Wang, MD[1]

[1]Departments of Spine Surgery, The First Hospital of Jilin University, Changchun, Jilin, China and [2]Nursing Platform of Spinal Surgery Department and [3]Department of Hospice, The First Hospital of Jilin University, Changchun, Jilin, People's Republic of China

Objective: This study aimed to build a predictive model of lower lumbar instability.

Methods: This retrospective study included 199 patients. Patients were divided into the lower lumbar instability group (LLIG) (n = 98) and lower lumbar stability group (LLSG) (n = 101). All participants of LLIG were recruited over a 2-year period (2015–2017) from the patients who accept lumbar surgery at the First Hospital of Jilin University. The LLSG was selected from outpatients who had underwent lumbar spine computed tomography (CT) and Flexion and extension radiographs (FER) at the First Hospital of Jilin University from 2015 to 2017. Several lower lumbar parameters were measured, including Lordosis angle (LA), intervertebral height (IH), ratio of anterior height to posterior height (APR), angle between endplate and anterior edge of vertebral body (AEPVa), sagittal slip ratio (SSR), and angle between the upper endplate and z-axis on sagittal plane (AUEZS). These parameters were keyed into the SPSS software to create a predictive model for classification. Sensitivity, specificity, predictive accuracy, and Kappa value were used to evaluate the predictive model.

Results: Compared with LLSG, the LA of LLIG decreased by 3.49° (126.54° *vs* 130.3°). Similarly, the IH of LLIG decreased by 1.23°mm, 1.66°mm, and 0.71°mm at L3-4, L4-5, and L5-S1. Compared with LLSG, the SSR of LLIG is higher at L3-4, L4-5, and L5-S1 (0.54 *vs* 0.51, 0.57 *vs* 0.46, and 0.59 *vs* 0. 47). Moreover, the APR of LLIG is higher than those of LLSG at L3-4, L4-5, and L5-S1 (1.97 *vs* 1.81, 2.40 *vs* 1.97, and 2.69 *vs* 2.26). The LLIG has bigger AEPVa than LLIG at L3-4, L4-5, and L5-S1. Compared with LLSG, the AUEZS of LLIG is bigger at L3-4 (91.75° *vs* 90.81°) and smaller at L4-5 and L5-S1(84.63° *vs* 85.85° and 73.27° *vs* 75.01°). The SSR (L4) show highest predictive accuracy (83%) when every parameter was fed to LDA classifier to generate a univariate model. All parameters represent a statistically significant difference ($P < 0.05$) between LLSG and LLIG. The model including LA, APR (L5-S1), IH (L4-5), SSR (L5), AUEZS (L5) has highest predictive accuracy of 88.2%. The sensitivity, specificity, and Kappa value are 88.7%, 93.1%, and 0.77.

Conclusion: The predictive model has good classification performance and can be an auxiliary tool for clinicians to evaluate lumbar instability in preoperative patients with severe pain aggravated by lumbar movement.

Key words: Three-dimensional computed tomographic; Lower lumbar instability; Predictive model

Address for Correspondence Zhenyu Wang MD, Department of Spine surgery, The First Hospital of Jilin University, No. 1, Xinmin St, Changchun, Jilin, 130021 China; Fax: 0431-80763563, Tel: +86 13578728836, Email: spinely@163.com
Zijing Liu and Yanbing Kao contributed equally to this work.
Received 29 April 2020; accepted 18 October 2020

Orthopaedic Surgery

Introduction

Lumbar spine instability is defined as abnormal lumbar motion during physiological loading of the spine[1–4]. Evaluating spinal instability is necessary to decide whether fusion is needed. Compared with decompression surgery alone, lumbar fusion surgery tends to have greater blood loss, longer hospital stay, and higher costs. Thus, patients should be adequately assessed for lumbar spine instability to avoid subjecting them to unnecessary fusion surgery[5].

Currently, there is no consensus regarding the criteria for diagnosing lower spine instability. Various imaging modalities have been used, including magnetic resonance imaging (MRI), CT, and FER. The most common method to assess lumbar instability is FER[6,7]. By performing FER in patients, the extent of vertebral translation and degree of angular displacement can be obtained, which can be used to determine the presence of abnormal segment movement. Although FER can reflect abnormal segment movement directly, this method is very difficult to apply in those patients who have severe back and leg pain aggravated by lumbar movement, and such patients are not rare in spine surgery departments. Thus, a comfortable and accurate method of assessing lumbar instability is necessary.

In the last decade, we have seen an increase in the number of articles devoted to the development of predictive models coming from the machine-learning community[8]. These models use a predefined set of relevant features but combine their values according to different classification paradigms, which, moreover, can automatically redefine the predictive models built when adding new records to the database. The classifiers usually obtain very good performance results after validation and are used as decision-support systems for clinician.

Lumbar instability was defined as a loss of lumbar stiffness, and such loss was related to damage to the restraining structures[4]. Panjabi *et al.* developed a checklist for the diagnosis of lumbar spine instability[9]. In this checklist, the anterior and posterior structure destroyed or unable to function were included, which indicated that lumbar spine instability is closely related to the anatomical changes. Leone *et al.* maintain that spinal degenerative changes including disc degeneration, facet joint osteoarthritis, and ligament degeneration play a role in the development of lumbar spine instability and degenerative spondylolisthesis[10]. We measured the radiological parameters which are related to the above anatomical changes and use them to build a predictive model.

Compared with MRI, we recognize that CT scans have the added advantage of three-dimensional visualization of the spinal instability assessment and a higher resolution for bony details[11]. The CT scan of the lumbar spine has been taken in the preoperative routine and the patient does not need to accept extra radiation exposure. Thus, we measured radiological parameters on three-dimensional computed tomography (3D-CT).

The purposes of this study were: (i) to explore the difference of lumbar anatomical parameters between LLSG and LLIG; (ii) to build a predictive model for preoperative patients; and (iii) to verify the accuracy of this predictive model and consistency between FER and predictive model.

TABLE 1 Demographic information of the 199 observations

	LLSG	LLIG	*P*-value
Numbers	98	101	
Gender	48 (male) 50 (female)	50 (male) 51 (female)	0.863
Age (years)	43.5 ± 8.6	45.6 ± 6.8	0.083
Height (cm)	170 ± 23	172 ± 31	0.190
Weight (kg)	72.4 ± 16	76.7 ± 23	0.157

Methods and Materials

This is a retrospective case–control pilot study. Before the commencement of the study, the approval of the Ethics Committee of First Hospital of Jilin University was sought, and written informed consent was obtained from all subjects.

Inclusion and Exclusion Criteria

LLIG inclusion criteria included: (i) all participants of LLIG were recruited over a 2-year period (2015–2017) from the patients who accepted lumbar surgery at the First Hospital of Jilin University patients; (ii) the preoperative patients were diagnosed as lumbar disc herniation with severe nerve root pain; and (iii) imaging performance: sagittal translation exceeds 4 mm or sagittal plane rotation exceeds cut-off (15° at L1–L2, L2–L3, and L3–L4; 20° at L4–L5; and 25° at L5–S1) on lumbar FER.

LLSG inclusion criteria included: (i) the outpatients who underwent lumbar spine CT and FER at the First Hospital of Jilin University from 2015 to 2017; (ii) the symptom of patients is low back pain without the being caused by lumbar nerve root compression; (iii) imaging performance: sagittal translation under 4 mm and sagittal plane rotation under cut-off (15° at L1–L2, L2–L3, and L3–L4; 20° at L4–L5; and 25° at L5–S1) on lumbar FER; and (iv) and the radiological report showed no lumbar abnormality.

Exclusion criteria included: (i) participants with endocrine system diseases (diabetes mellitus, hyperthyroidism, hypothyroidism, Hashimoto's thyroiditis, hyperparathyroidism, and hypoparathyroidism); (ii) any history of spinal trauma (vertebral compression fractures and vertebral burst fractures); (iii) previous history of spinal surgery (lumbar decompression surgery, lumbar spinal fusion surgery, and other surgical procedures that may affect the anatomy of the lumbar spine); (iv) suspected findings of spinal malignancies (abdominal neoplasms and spinal tumor); (v) any history of inflammatory spinal diseases (spinal brucellosis infection, spinal tuberculosis, intervertebral infection, and vertebral infection); and (vi) congenital and developmental anomalies (idiopathic scoliosis and congenital spine structure abnormality).

Demographic Characteristic

Each patient's relevant medical history was recorded as well as patient details including age, sex, height, and weight. The demographic characteristics of these subjects are listed in Table 1. Lumbar instability was present on different segments in LLIG with the level L3-4 in three, L4-5 in 69, and L5-S1 in 29 cases.

Measurement Methods and Parameters

The sagittal, coronal, and axial view of CT scan was reconstructed in workstation (Advantage Workstation 4.3; GE Medical Systems). The reconstructed 3D-CT models could be rotated, cut, clipped, and measured. Six parameters were measured on 3D-CT model. Measurements were made independently by two blinded authors and averaged to ensure accuracy, and inter-observer reliability was excellent (Pearson's $P = 0.821$).

The definition and clinical significance of measured parameters were listed below:

Lordosis Angle (LA)

Definition. The lordosis angle of lumbar spine.

Measurement Method. On central sagittal plane, we set lower right point of T12 as A point, upper right point of S1 as B point, and midpoint of L3 front edge as O point. ∠AOB is LA (Fig. 1).

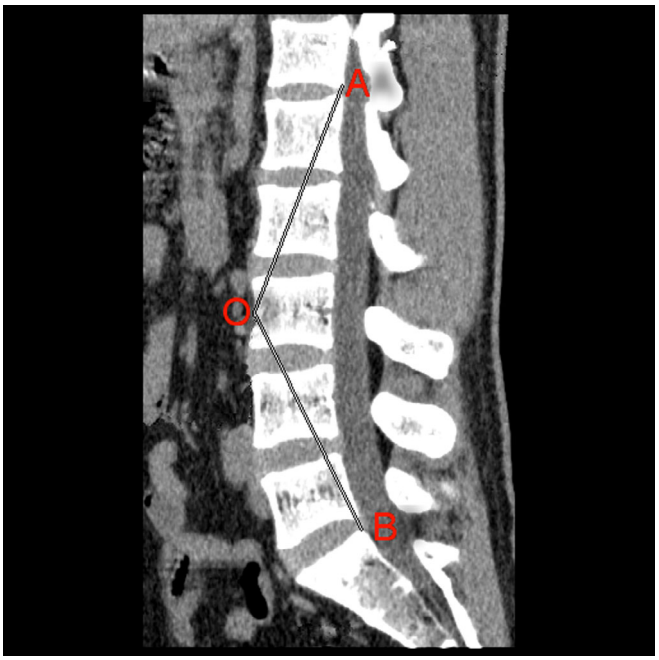

Fig 1 The lordosis angle (LA) measured on sagittal plane.

Clinical Significance. This parameter is an indicator of height change of whole lower lumbar spine. The bigger this angle, the more the disc height is decreased.

Intervertebral Height (IH)

Definition. The height of intervertebral space between upper and lower endplate.

Measurement Method. The distance from midpoint of lower endplate of superior vertebra to midpoint of upper endplate of inferior vertebra on central sagittal plane of 3D-CT (Fig. 2).

Clinical Significance. Decreased intervertebral height is one of most important factors which lead to lower lumbar instability.

Ratio of Anterior Height to Posterior Height (APR)

Definition. The ratio of anterior IH to posterior IH (Fig. 3).

Measurement Method. On central sagittal plane of 3D-CT, we define the distance from anterior edge of lower endplate of superior vertebra to anterior edge of upper endplate of inferior vertebra as anterior IH. Similarly, we measured posterior IH by using the same method.

Clinical Significance. When the height of anterior and posterior intervertebral spaces decrease asymmetrically, the sagittal rotation could happen. This parameter can be used to evaluate the degree of vertebral rotation which may lead to vertebral abnormal movement.

Angle Between Endplate and Anterior Edge of Vertebral Body (AEPVa)

Definition. The angle between upper endplate of inferior vertebra and anterior edge of superior vertebra (Fig. 4).

Measurement Method. On central sagittal plane of 3D-CT, we determined angle between upper endplate of inferior vertebra and anterior edge of superior vertebra.

Clinical Significance. This parameter was used to evaluate the degree of sagittal vertebral rotation based on inferior vertebra. This situation is related to the reduction of disc height and laxity of the ligament.

Coordinate System

A standard coordinate system was established on 3D-CT model. The midpoint of the sacrum's anterior edge was set as point O (Fig. 5).

Sagittal Slip Ratio (SSR)

Definition. The ratio of distance from anterior point of upper endplate to the length of upper endplate.

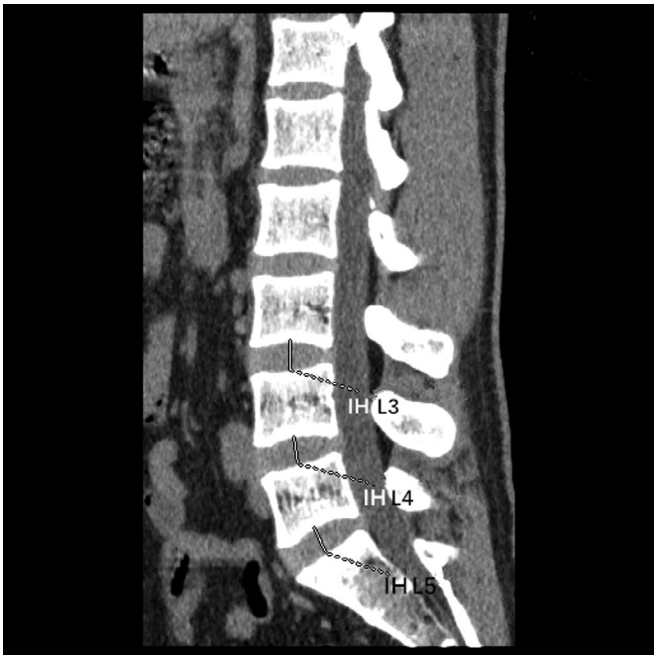

Fig 2 The intervertebral height (IH) measured on sagittal plane.

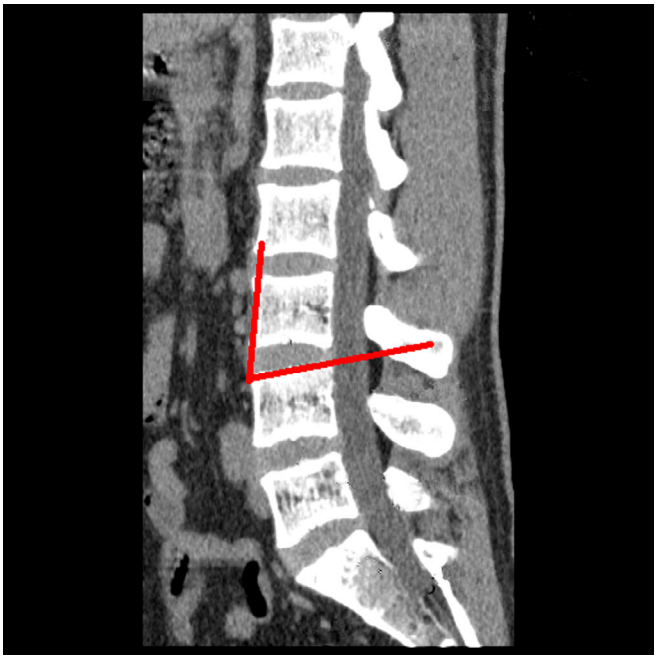

Fig 4 The AEPVa measured on sagittal plane.

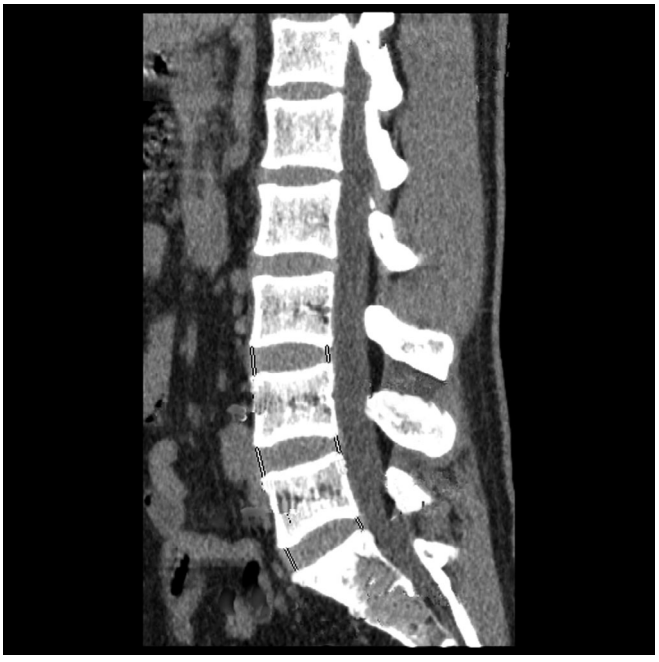

Fig 3 The anterior height and posterior height of disk measured on sagittal plane.

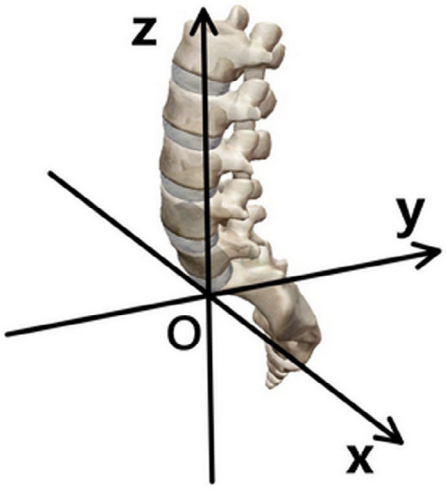

Fig 5 Three-dimensional coordinate system. O, midpoint of the front edge of sacrum. The x-, y-, and z-axes are vertical to each other as intersecting lines of the sagittal, horizontal, and coronal planes.

Measurement Method. On the sagittal plane including z-axis, the distance from the anterior edge of upper endplate to z-axis was defined as DAUEZ. Similarly, the distance from the posterior edge of upper endplate to z-axis was defined as DPUEZ. SSR = **DAUEZ/** (**DAUEZ + DPUEZ**) (Fig. 6).

Clinical Significance. This parameter is an indicator of the degree of sagittal slip. Measurement in this manner ensured adjustment for any differences in magnification, by standardizing against the length of upper endplate.

Angle Between the Upper Endplate and Z-Axis on Sagittal Plane (AUEZS)

Definition. The angle between right half upper endplate and z-axis of coordinate system (Fig. 7)

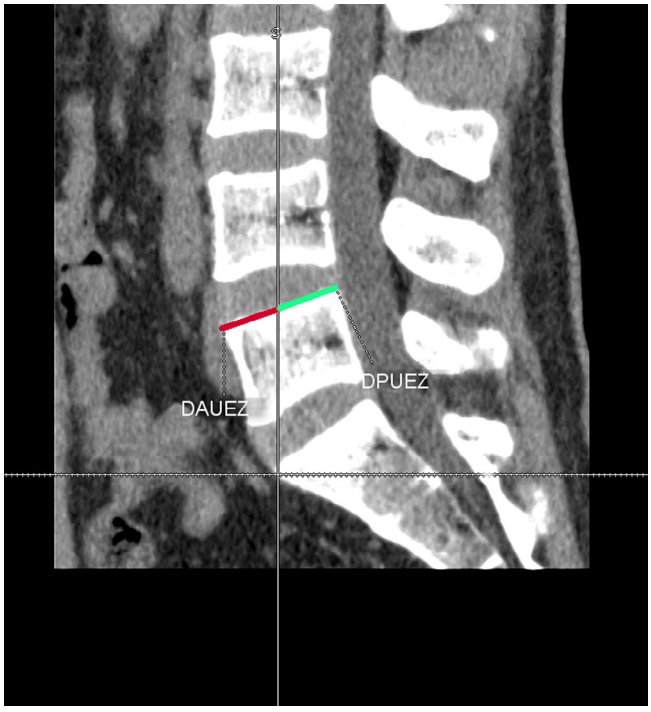

Fig 6 The DAUEZ and DPUEZ measured on sagittal plane including z-axis.

Measurement Method. On central coronal plane of 3D-CT, we determined the angle between right half upper endplate and z-axis of coordinate system.

Clinical Significance. This parameter is an indicator of sagittal imbalance caused by decrease in disc height and acquired collapse.

Data Analysis

Linear discriminant analysis (LDA) develops a discriminant criterion and then finds an optimal discriminant function to efficiently classify each participant into one of the groups. Because of its simplicity and effectiveness, it has been widely used in many fields. To avoid the curse of dimensionality (i.e. as dimensions increase, the volume of space increases exponentially and the available data becomes sparse), a parameter selection method called stepwise discriminant analysis was used. By applying stepwise discriminant analysis, the best subset of parameters can be determined. We then fed the selected parameter to the linear discriminant classifier. The predictive model of lower lumbar instability was created. To measure the performance of the classification model, we used sensitivity, specificity, and predictive accuracy, which were computed based on the leave-one-out cross-validation (LOOCV).

The commonly used measure of a model's quality is predictive accuracy. The predictive accuracy is usually estimated from all the available samples, which are split into training and testing sets. We understand that the training and the test sets should be of sufficient size and independent in character so that it can be used to generate a reliable estimate of predictive accuracy. We had two important issues to consider when we chose a method to estimate predictive accuracy. There was concern that our model will not be very robust because the training set was small and, consequently, we may not be able to generalize the findings to other independent datasets. However, the confidence in the estimated predictive accuracy will be low because our test set was also small. We therefore chose a method to specifically deal with the limitations in our data. In our study, we used LOOCV to get over the limitation of data. This method is designed to test predictive accuracy within a single small dataset. It uses a single observation from the original sample as the validation data and the remaining observations are the training data. Thus, we used (n − 1) samples for training and evaluated the remaining sample. This is repeated so that each observation in the sample is used once as the validation data. This is a well-accepted approach but is used less often because it has large computational requirements. There are many examples of LOOCV in the medical field[12, 13].

In this study, we aimed to build a predictive model based on the data of 199 observations. The parameter data of 199 observations were input into the SPSS. Every parameter was fed to LDA classifier to generate an univariate model. Then we get the average, standard deviation, *P*-value, and predictive accuracy based on LOOCV of every parameter by SPSS. Finally, we fed the whole parameters into LDA classifier to create a multivariate predictive model. The sensitivity, specificity,

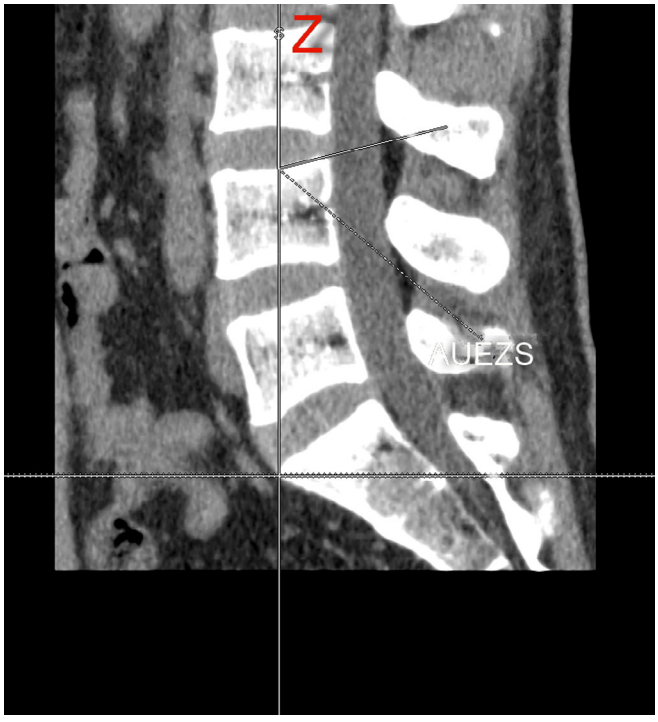

Fig 7 The AUEZS measured on sagittal plane including z-axis.

TABLE 2 The average, standard deviation, *P*-value, classified accuracy, and Kappa value of univariate model calculated by SPSS

	LLSG		LLIG		*P*	LOOCV (%)	Kappa value
	Means	SD	Means	SD			
LA	130.03	6.10	126.54	6.34	0.001	82.3	0.64
APR (L3-4)	1.81	0.56	1.97	0.62	0.016	54.5	0.32
APR (L4-5)	1.97	0.54	2.40	1.84	0.002	57.9	0.29
APR (L5-S1)	2.26	0.68	2.69	0.78	0.037	62.8	0.45
IH (L3-4)	10.42	1.79	9.19	2.09	0.046	80.7	0.61
IH (L4-5)	10.01	1.94	8.35	1.99	0.003	81.4	0.64
IH (L5-S1)	9.11	2.08	8.40	2.38	0.034	82.2	0.59
AEPVa (L3-4)	78.86	3.06	77.96	3.57	0.002	60.0	0.38
AEPVa (L4-5)	73.18	3.78	74.44	3.38	0.009	59.3	0.41
AEPVa (L5-S1)	67.19	4.34	67.73	4.56	0.004	54.5	0.46
SSR (L3)	0.51	0.64	0.54	0.71	0.048	80.2	0.63
SSR (L4)	0.46	0.52	0.57	0.43	0.002	83.0	0.65
SSR (L5)	0.47	0.33	0.59	0.28	0.004	81.2	0.61
AUEZS (L3)	90.81	5.06	91.75	6.26	0.036	57.2	0.37
AUEZS (L4)	85.85	5.23	84.63	5.86	0.007	57.2	0.34
AUEZS (L5)	75.01	5.90	73.27	5.96	0.009	57.9	0.31

LOOCV, leave-one-out cross validation; SD, standard deviation.

predictive accuracy, and Kappa value was calculated. The whole process of LDA and LOOCV was completed by SPSS.

Results

The Demographic Characteristic

Five demographic parameters were collected including number of patients in each group, gender, age, height, and weight. These parameters were compared between the two groups by SPSS software. The mean value, standard deviation, and *P* value calculated by *t* test were listed below: Age (43.5 ± 8.6 *vs* 45.6 ± 6.8, *P* value = 0.083), gender (48 male and 50 female *vs* 50 male and 51 female, *P* value =0.863), height (170 ± 23 cm *vs* 172 ± 31 cm, *P* value = 0.190) and weight (72.4 ± 16 kg *vs* 76.7 ± 23kg, *P* value = 0.157).

P values of all parameters were higher than 0.05, which indicates no significant difference in all parameters studied.

Lumbar Anatomical Parameters Change with Lower Lumbar Instability

To build a predictive model for evaluating the lower lumbar instability, we collected CT data of 199 patients and measured six radiological parameters on reconstructed 3D-CT images. The mean value, standard deviation, *P* value calculated by *t* test, the accuracy calculated by LOOCV and Kappa value were listed in Table 2. As demonstrated in Table 2, *P* values of all parameters were less than 0.05, which indicates significant differences in all parameters studied between two groups.

Compared with LLSG, the LA of LLIG decreased by 3.49° (126.54° ± 6.34° *vs* 130.3° ± 6.10°). The difference indicates that the lordosis of lumbar spine increases when lower lumbar instability occurs.

The IH was measured at L3-4, L4-5, and L5-S1 levels. According to Table 2, the average IH of LLIG is 9.19 ± 2.09 mm, 8.35 ± 1.99 mm, and 8.40 ± 2.38 mm. The average IH of LLSG is 10.42 ± 1.79 mm, 10.01 ± 1.94 mm, and 9.11 ± 2.08 mm. Compared with LLSG, the IH of LLIG is lower, especially at L4-5 level. It suggests that the reduction of intervertebral height may be related to the occurrence of lumbar instability.

Moreover, APR and AEPVa were measured on 3D-CT at L3-4, L4-5, and L5-S1 levels. The APR of LLIG is higher than those of LLSG at L3-4, L4-5, and L5-S1 levels (1.97 *vs* 1.81, 2.40 *vs* 1.97, and 2.69 *vs* 2.26). The AEPVa of LLIG is 77.96° ± 3.57°, 74.44° ± 3.38°, and 67.73° ± 4.56°. LLSG has smaller AEPVa at L3-4,L4-5, and L5-S1 levels (78.86° ± 3.06°, 73.18° ± 3.78°and 67.19° ± 4.34°). Compared with LLSG, AEPVa increase by 1.26° and 0.54° at L3-4 and L5-S1 levels, which indicates the occurrence of vertebral rotation. These results show that lower lumbar instability may be related to vertebral rotation caused by reduction of disc height and laxity of the ligament.

As demonstrated in Table 2, we measured SSR on established coordinate system. The SSR of LLIG is 0.54, 0.57, and 0.59. The SSR of LLSG is 0.51, 0.46, and 0.47. Compared with LLSG, SSR of LLIG increased by 0.03, 0.11, and 0.12. The result indicates that lower lumbar instability may happen with more severe sagittal slip.

According to Table 2, AUEZS of LLIG is 91.75°, 84.63°, and 73.27° at L3-4, L4-5, and L5-S1 levels. The AUEZS of LLSG is 90.81°, 85.85°, and 75.01° at L3-4, L4-5, and L5-S1 levels. We can know that AUEZS of LLIG is bigger at L3-4 (91.75° *vs* 90.81°) and decreased by 1.22° and 1.74° at L4-5 and L5-S1 (84.63° *vs* 85.85°and 73.27° *vs* 75.01°) compared with LLSG.

TABLE 3 Test results of the prediction model

	LLIG	LLSG
Positive	86	7
Negative	12	94

The Results of Predictive Accuracy and Kappa Value of Univariate Models

Every parameter was fed to LDA classifier to generate an univariate model. The accuracy of univariate model calculated by LOOCV is presented comparatively in Table 2. As demonstrated in Table 2, seven parameters have predictive accuracy higher than 80%, including LA, IH (L3-4), IH (L4-5), IH (L5-S1), SSR (L3), SSR (L4), and SSR (L5). The SSR (L4) show the highest predictive accuracy (83%), which indicates it is a very useful indicator to diagnose the lower lumbar instability. Also, the predictive accuracy of other parameters is lower than or equal to 62.8%, which indicates these parameters have poor predictive performance. As demonstrated in Table 2, the Kappa value of LA, IH (L3-4), IH (L4-5), SSR (L3), SSR (L4), and SSR (L5) was higher than 0.6, which indicates that these parameters have satisfactory consistency.

Predictive Accuracy, Sensitivity, Specificity, and Kappa Value of Multivariate Predictive Model

We used machine-learning methods to identify subsets of radiological parameters with the highest predictive accuracy. The LDA was applied to all possible subsets derived from the 16 radiological parameters. The subset consisting of LA, APR (L5-S1), IH (L4-5), SSR (L5), and AUEZS (L5) has highest predictive accuracy of 88.2%. This subset was defined as set A.

The origin data was tested by a predictive model created by set A. The result was listed in Table 3. The sensitivity and specificity are 88.7% and 93.1%. The Kappa value was used to evaluate the consistency between FER and predictive model. The Kappa value is 0.77, which indicates a good level of consistency.

The Result of Subgroup Analyses

We collected five demographic parameters – number of patients in each group, gender, age, height, and weight. To the best of our knowledge, previous studies have revealed the existence of lumbar anatomical difference between males and females. We assumed that the anatomical difference related to the gender may aggravate when lumbar instability occurred. Then, we carried out subgroup analyses to verify this hypothesis.

According to the gender, the LLIG is divided into two group: LLIG (male) 48 and LLIG (female) 50. Similarly, the LLSG is also divided into LLSG (male) 50 and LLSG (female) 51. Then, we perform sexual subgroup analysis. The results of separate analyses for females and males are presented in the Tables 4–7. As demonstrated in Tables 4, 5, the male subgroup analysis gets similar results as the analysis without considering gender. The accuracy of multivariate predictive model created by stepwise discriminant analysis is 88.9% (*vs* 88.2%). The sensitivity, specificity, and Kappa value are 89.5% (*vs* 88.7%), 90.0% (*vs* 93.1%), and 0.78 (*vs* 0.77).

And the result of female subgroup analysis is presented in Tables 6, 7. The accuracy of multivariate predictive model created by female subgroup analysis is 88.1% (*vs* 88.2%). The sensitivity, specificity, and Kappa value are 88.0% (*vs* 88.7%), 94.1% (*vs* 93.1%), and 0.75 (*vs* 0.77). The result shows there is no significant difference between whole population analysis and female subgroup analysis.

In conclusion, it is observed that the results of subgroup analyses (for gender) are consistent with the results of whole population analysis.

Discussion

Current Research Status of the Evaluation of Lower Lumbar Instability

Since the concept of lumbar instability was introduced in the 1980s, the definitions of lumbar instability have always been controversial. However, a reasonable definition has been proposed by Pope and Panjabi[4], and Frymoyer and Selby[2].They advocated a biomechanical approach and defined instability as a loss-of-motion segment stiffness, such that force application to that motion segment produces abnormal gross movements compared to that of a normal spine. This abnormal movement can be explained by the damage to restraining structures, such as the facet joints, discs, ligaments, and muscles; if these structures are damaged or became lax, altered equilibrium and thus instability occur[4]. When nerve symptom arises from abnormal movement of the spine, decompression often cannot be adequately achieved unless the spine is stabilized and fused.

The most common diagnostic method is FER, which evaluates lower lumbar instability using angular change and vertebral displacement[14]. Although FER emits low radiation and is inexpensive, limited range of motion caused by pain-related fear misguides the diagnosis and surgical decision making. Many clinical doctors realize this problem and make efforts to improve it. For example, Hey *et al.* improved the FER using the slump sitting posture to replace the standing posture in flexion. The new method of slump sitting dynamic radiography was shown to have higher efficacy than FER in measuring the angular change. However, this method still requires patients to endure pain in executing complete flexion and extension movements as much as possible. It is still very painful and difficult for some pain-sensitive patients.

The Advantage of Predictive Model on Evaluation of Lower Lumbar Instability

In this study, we innovatively used the method of establishing a coordinate system to assess the slip degree of the vertebral body. Since the sacrum is connected to the pelvis, its position can barely change. Thus, we set a standard coordinate system, with the midpoint of the sacrum's front edge

TABLE 4 The average, standard deviation, *P*-value, classified accuracy, and Kappa value of male subgroup analysis calculated by SPSS

	LLSG (male)		LLIG (male)		*P*	LOOCV (%)	Kappa value
	Means	SD	Means	SD			
LA	131.04	5.91	127.56	6.54	0.011	83.3	0.66
APR (L3-4)	1.83	0.61	1.98	0.72	0.035	56.5	0.37
APR (L4-5)	2.01	0.84	2.73	1.85	0.001	58.9	0.39
APR (L5-S1)	2.28	1.52	2.74	1.78	0.027	61.7	0.41
IH (L3-4)	10.46	3.79	9.21	4.09	0.016	80.9	0.58
IH (L4-5)	10.04	1.13	8.39	1.36	0.001	84.4	0.67
IH (L5-S1)	9.17	2.65	8.42	3.38	0.048	82.1	0.58
AEPVa (L3-4)	78.96	3.26	78.06	4.57	0.001	61.7	0.48
AEPVa (L4-5)	73.68	3.98	74.94	5.38	0.019	59.9	0.40
AEPVa (L5-S1)	67.59	3.89	67.96	5.57	0.014	57.5	0.46
SSR (L3)	0.52	1.23	0.54	1.71	0.028	80.7	0.68
SSR (L4)	0.45	1.52	0.59	2.43	0.032	82.0	0.63
SSR (L5)	0.49	1.33	0.62	0.98	0.001	81.7	0.62
AUEZS (L3)	90.99	4.06	91.83	4.26	0.002	58.2	0.41
AUEZS (L4)	85.96	2.23	84.78	5.73	0.017	57.5	0.37
AUEZS (L5)	75.12	3.18	73.64	2.96	0.003	58.9	0.34

LOOCV, leave-one-out cross validation; SD, standard deviation.

TABLE 5 Test results of prediction model of male subgroup analysis

	LLIG (male)	LLSG (male)
Positive	43	5
Negative	5	45

as point O. In this system, SSR show the degree of lumbar vertebral sagittal displacement. The anterolisthesis usually happens along with severe disc degeneration, facet joint osteoarthritis, and ligament degeneration. Similarly, we can determine the degree of lumbar vertebral rotation from AUEZS.

Moreover, we analyzed six radiological parameters acquired from 3D-CT, all of which had obvious statistical differences ($P < 0.05$). The univariate model (LA, SSR, and IH) has predictive accuracy which is higher than 80%. The predictive accuracy of a univariate model reached up to 80%, indicating that the parameter has significant correlation with lumbar spine instability. Furthermore, the three parameters mainly reflect the sagittal displacement and intervertebral

TABLE 6 The average, standard deviation, *P*-value, classified accuracy, and Kappa value of female subgroup analysis calculated by SPSS

	LLSG (Female)		LLIG (Female)		*P*	LOOCV (%)	Kappa value
	Means	SD	Means	SD			
LA	129.02	6.83	125.54	7.54	0.021	80.3	0.63
APR (L3-4)	1.79	0.98	1.96	1.72	0.016	57.5	0.38
APR (L4-5)	1.93	1.84	2.08	3.85	0.007	58.2	0.39
APR (L5-S1)	2.24	2.86	2.64	1.49	0.041	64.8	0.43
IH (L3-4)	10.38	1.89	9.17	2.06	0.036	81.7	0.62
IH (L4-5)	10.04	4.52	8.31	3.95	0.001	81.9	0.65
IH (L5-S1)	9.05	2.78	8.38	4.38	0.004	81.1	0.62
AEPVa (L3-4)	78.76	3.96	78.86	4.57	0.009	60.0	0.45
AEPVa (L4-5)	72.70	5.78	73.94	4.38	0.017	59.4	0.42
AEPVa (L5-S1)	66.80	6.81	67.50	5.56	0.005	54.9	0.45
SSR (L3)	0.50	2.67	0.54	2.71	0.018	80.6	0.64
SSR (L4)	0.47	1.52	0.55	2.43	0.022	82.8	0.66
SSR (L5)	0.45	1.33	0.56	2.28	0.014	81.8	0.62
AUEZS (L3)	90.63	5.16	91.67	7.26	0.018	57.4	0.39
AUEZS (L4)	85.74	4.23	84.48	5.17	0.003	56.2	0.33
AUEZS (L5)	74.90	3.90	72.90	4.83	0.004	57.8	0.32

LOOCV: leave-one-out cross validation; SD, standard deviation.

TABLE 7 Test results of prediction model of female subgroup analysis

	LLIG (female)	LLSG (female)
Positive	44	3
Negative	6	48

disc height change of the vertebral body. Several biomechanical and clinical studies have reported the association of disc degeneration with segmental instability[2, 9, 10, 15, 16]. They reported that an anterior translation of 3 mm or greater was positively associated with disc degeneration and facet joint osteoarthritis. Thus, our result is consistent with that of a previous study. The multivariate predictive model tested using original data demonstrated sensitivity and specificity of 88.7% and 93.1%, respectively. This shows that our predictive model has good classification performance. The Kappa value is 0.77, which indicates a good level of consistency.

Limitation of the Study

Nevertheless, our proposed method does have some limitations. Firstly, this is a retrospective study, and a prospective study should be performed to exclude the effect of other factors. Secondly, the mechanism by which lumbar anatomical parameters change wasn't studied in this research. It is our goal to explore the mechanism in the future.

Conclusion

The predictive model has good classification performance and can be an auxiliary tool for clinicians to evaluate lumbar instability.

References

1. Dupuis PR, Yong-Hing K, Cassidy JD, Kirkaldy-Willis WH. Radiologic diagnosis of degenerative lumbar spinal instability. Spine, 1985, 10: 262–276.
2. Frymoyer JW, Selby DK. Segmental instability. Rationale for treatment. Spine, 1985, 10: 280–286.
3. Izzo R, Guarnieri G, Guglielmi G, Muto M. Biomechanics of the spine. Part I: spinal stability. Eur J Radiol, 2013, 82: 118–126.
4. Pope MH, Panjabi M. Biomechanical definitions of spinal instability. Spine, 1985, 10: 255–256.
5. Hey HW, Lau ET, Lim JL, *et al*. Slump sitting X-ray of the lumbar spine is superior to the conventional flexion view in assessing lumbar spine instability. Spine J, 2017, 17: 360–368.
6. Boden SD, Wiesel SW. Lumbosacral segmental motion in normal individuals. Have we been measuring instability properly? Spine, 1990, 15: 571–576.
7. Nachemson AL. Instability of the lumbar spine. Pathology, treatment, and clinical evaluation. Neurosurg Clin N Am, 1991, 2: 785–790.
8. Chun FK, Karakiewicz PI, Briganti A, *et al*. A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer. BJU Int, 2007, 99: 794–800.
9. Fujiwara A, Lim TH, An HS, *et al*. The effect of disc degeneration and facet joint osteoarthritis on the segmental flexibility of the lumbar spine. Spine, 2000, 25: 3036–3044.
10. Leone A, Guglielmi G, Cassar-Pullicino VN, Bonomo L. Lumbar intervertebral instability: a review. Radiology, 2007, 245: 62–77.
11. Larde D, Mathieu D, Frija J, Gaston A, Vasile N. Spinal vacuum phenomenon: CT diagnosis and significance. J Comput Assist Tomogr, 1982, 6: 671–676.
12. Cheng H, Garrick DJ, Fernando RL. Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction. J Anim Sci Biotechnol, 2017, 8: 38.
13. Seitsonen ER, Korhonen IK, van Gils MJ, *et al*. EEG spectral entropy, heart rate, photoplethysmography and motor responses to skin incision during sevoflurane anaesthesia. Acta Anaesthesiol Scand, 2005, 49: 284–292.
14. Panjabi MM. Clinical spinal instability and low back pain. J Electromyogr Kinesiol, 2003, 13: 371–379.
15. Kazarian, L. Dynamic Response Characteristics of the Human Vertebral Column. An Experimental Study on Human Autopsy Specimens. Acta Orthopaedica Scandinavica. Supplementum. 1972: 1–186.
16. Weiler PJ, King GJ, Gertzbein SD. Analysis of sagittal plane instability of the lumbar spine in vivo. Spine, 1990, 15: 1300–1306.