Research Article

Open Access

Reliability and Validity of the Side-lying Instability and Prone Instability Tests in Patients with Lumbar Segmental Instability

Bo—Eon Kim, PT, MS · Kwan—Woo Lee, PT, PhD[1] · Dae—Sung Park, PT, PhD[2][†]

Graduate School of Public Health & Welfare, Konyang University,
[1]Department of Physical Therapy, Ulsan College,
[2]Department of Physical Therapy, Konyang University

| Abstract |

PURPOSE: The purpose of this study is to conduct inter-rater and intra-rater reliability tests in patients with low back pain (LBP) using the prone instability test (PIT) and side-lying instability test (SIT). We have analyzed the Korean version Oswestry disability index (K-ODI) correlations and radiograph finding (RF) for validity.

METHODS: Individuals (n = 51) (mean age of 40.27 ± 13.28) with LBP for at least over a week were recruited, together with two participating physical therapist examiners. The measurement consisted of PIT, PST, K-ODI, and RF. Sensitivity (Sn), specificity (Sp), positive predictive value, negative predictive value, prevalence index, agreement %, Cohen's kappa, and prevalence-adjusted bias-adjusted kappa (PABAK) were calculated. The PIT and SIT were compared with RF for validity analysis, while PIT, SIT, K-ODI, and RF were calculated for the correlation analysis.

RESULTS: The intra-rater reliability test measured for the PIT (kappa = .79, PABAK = .88) and SIT (kappa = .73, PABAK = .84), and inter-rater reliability test measured for the SIT (kappa = .80, PABAK = .88) showed good agreements. The PIT (Sn = .65, Sp = .63) and SIT validities (Sn = .68, Sp = .70) were compared with RF, showing a significant correlation in PIT and RF (r = .69), SIT and RF (r = .73), and PIT and K-ODI (r = .53).

CONCLUSION: The SIT is a more comfortable position test than the PIT in patients. Both PIT and SIT have acceptable reliability and validity.

Key Words: Low back pains, Joint instability, Lumbar stenosis, Lumbosacral plexus, Instability test

†Corresponding Author : Dae-Sung Park
daeric@konyang.ac.kr, https://orcid.org/0000-0003-4258-0878

Ⅰ. Introduction

The physical therapists' clinical assessment ability and specialized clinical reasoning play an essential role in planning for low back pain (LBP) and functional recovery as they evaluate patients; conduct observation, palpation, and evaluation of joint mobility; and perform assistive exercise and muscle performance ability examination. To successfully treat LBP patients, physical therapists need

to classify further similar pain features and disabilities and build detailed treatment plans accordingly [1]. A loss of resistance to anterior shear force in the lumbar region has a possibility of tissue damage.

Joint degeneration causes excessive translation, while hypermobility angulation of the spinal units causes instability [2]. Diagnosis confirmation based on any single test between the general and special examinations is less reliable compared to the diagonal decision made based on combined multiple test results [3]. To diagnose LBP, mechanical processes, such as X-ray, CT, and MRI, are changeable. Physical therapists' assessment can be more accurate than mechanical methods, which can change quickly [4]. The active and passive movement-based LBP classification methods include the lumbar segmental instability (LSI) test [5,6], instability catch sign, and aberrant motions test [7]. LSI test assesses the extent of passive areas, such as the ligament, disk, and facet on the lumbar segment stabilization. Patients are instructed to relax as much as possible while passive test is performed for safety [3], which includes prone instability test (PIT) [8], posterior shear test [9, 10], etc. The PIT requires a person to be in prone position on the bed with the trunk supported as the therapist applies a passive anterior shear force over the lumbar spinous process. The test is recorded as a positive sign if a patient complains of any pain on one or more lumbar vertebral levels. However, the prone position for the PIT is uncomfortable for patients with LBP. Thus, we suggest the side-lying instability test (SIT), which is a modified PIT to examine lumbar instability in a more comfortable position for the patient.

The purpose of this study is to conduct inter-rater and intra-rater reliability tests for LBP patients using the PIT and SIT and a validity test by analyzing Cohen's Kappa, prevalence-adjusted bias-adjusted kappa (PABAK), the correlations in the Korean version Oswestry disability index (K-ODI) [11,12], and the radiograph finding (RF) [13].

II. Methods

1. Subjects

This study examined a prospective cohort of patients with LBP. Fifty-three patients (40.27 ± 13.28) with LBP participated, who are members of a local community in the Ulsan Metropolitan City in South Korea. Inclusion criteria were a chief LBP complaint without symptoms below the knee, occurring at least over a week, or reoccurred within the recent 6 months. The subjects were excluded if they had acute fracture, tumor, or inflammation problems; spinal disk herniation and fusion, weakened lower limb, radiating pain under knees, or loss of reflex; and are currently pregnant or with spinal deformation, neurological diseases, or other serious medical problems.

2. Procedures

Two physical therapists participated as raters and educated and consulted the necessary test technique in ten rounds of meetings for 1 hour each. Rater "A" was a physical therapist with a 12-year career in treating patients with musculoskeletal and central nervous system problems, while rater "B" was a physical therapist with a 9-year career in treating patients with musculoskeletal problems. The two physical therapists received an orthopedic manual therapy Kaltenborn-Evjenth concept spine course.

The subjects participated in both clinical measurements and the RF, which consisted of PIT, SIT, and K-ODI blind to the radiographic result. The first rater (A) performed all the clinical measurement measures (PIT, SIT, K-ODI, and RF) on each subject, while the second rater (B), blinded to the first evaluation results, performed PIT and SIT. The participants received an RF examination at a hospital in Ulsan (Fig. 1).

The PIT is used to confirm the instability of the lumbar segment from L1 to S1. The subjects take a prone position with a pillow placed between their pelvis and the bed, preventing an excessive lordotic curve and creating a
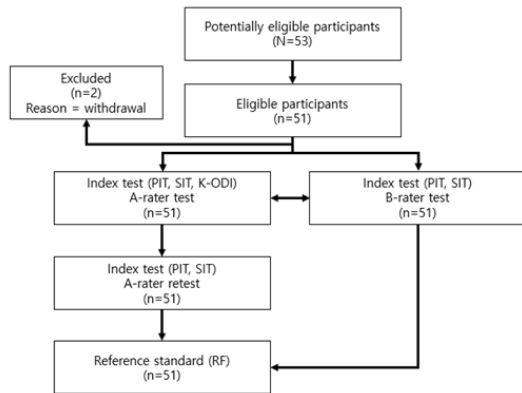
Fig. 1. Experimental flow diagram (K-ODI = Korean
version oswestry disability index, PIT = Prone
Instability Test, Side-lying Instability Test = SIT,
RF = Radiograph Finding).

Fig. 2. Prone instability test for Lumbar Segmental
Instability on L3 – L4.

comfortable position. The examiners find the upper-most
part of the subjects' iliac crest and draw a virtual line parallel
with the waist (Fig. 2). Force application to the vertebral
body's spinous process is a way of finding out if the
movements, like hypermobility and hypomobility, result
in the physiological range. If the sacrum is under L5, they
pressed on it frontward with the thenar eminence. If the
part gives a hard feeling, such as bone-to-bone, it can be
confirmed as S1, where the sacrum is located. Once the
S1 point is found, the examiners pressed frontward one
by one from L5 to L3 incrementally through the spinous
process. If a patient has any provocation of pain, it was
recorded as a positive [8].

Fig. 3. Side-lying instability test for Lumbar Segmental
Instability with an algometer on L4.

The SIT was assessed with the patient side lying on
the bed from L1 to S1. The therapists flex the patient's
hip joint to its maximum extent to distance the lumbar
segments, facilitating the ligament. For the spinous process
pressure consistency, place an electronic algometer (JTECH
Medical, USA) on the inter-spinous ligaments and perform
palpation between the segments. A positive sign occurred
if familiar symptoms were provoked, proposing a lumbar
instability indication (Fig. 3) [14].

The K-ODI includes ten elements with a score of 0
to 5: standing, sitting, walking, lifting, traveling, self-care,
employing/home activities, social participation, pain
intensity, and sleep. The total score of the K-ODI range
is 0 to 50, with high scores suggesting a high disability.
Jeon et al. [15] developed a modified Korean version of
ODI (K-ODI). The K-ODI's and test-retest's reliabilities
were Cronbach's alpha equal to 0.92 and 0.93, respectively.

Lumbar instability is tested as a sagittal plane
displacement from the lumbar spine's lateral radiographic
view at five levels (L1-L2, L2-L3, L3-L4, L4-L5, L5-S1).
Instability is positively recognized when the sagittal plane
displacement is at least 4.5 mm [2, 3, 4]. The magnitude
of the translation at each segment was calculated by
measuring the displacement of the superior vertebra and a
posterior vertical line of vertebral body of the inferior
vertebra. The displacement calculation on the radicular image
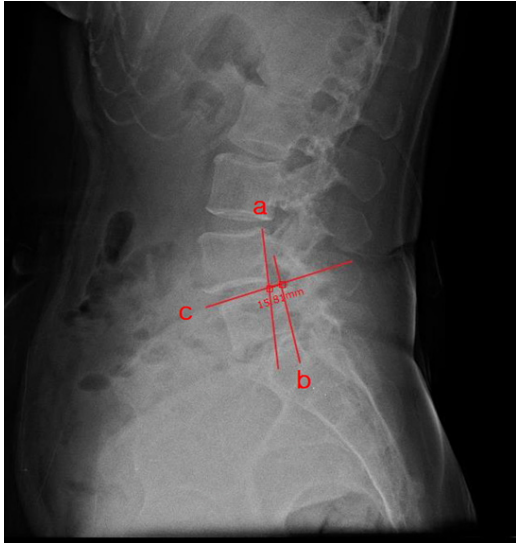was performed using the Sante Dicom image software

Fig. 4. A sample image of the sagittal plane displacement in radiograph finding on L4 - L5 (a: vertical line of vertebral body on L4, b: vertical line of vertebral body on L5, c: displacement distance between L4 and L5).

version 1.1.2 (Satesoft, Athens, Greece) (Fig. 4).

3. Statistics

The intra-rater and inter-rater reliability tests were used to analyze the reliability of the PIT and SIT, which are calculated by sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), prevalence ratio, agreement %, Cohen's kappa, and PABAK [18, 19]. The kappa values exist from no agreement (kappa = 0) to excellent (kappa = 1). The PABAK defines an index of agreement between two observers adjusting kappa for bias and the PABAK was calculated to provide the statistical adjustment of kappa given the prevalence and bias associated with testing. The kappa and PABAK values indexing reliability were interpreted as follows: less than .00 is poor; .00 to .20 is slight; .21 to .40 is fair; .41 to .60 is moderate; .61 to .80 is substantial; and .81 to 1.00 is almost perfect [8]. PABAK is zero when the observed agreement is equal to 50%. For the diagnostic

Table 1. General Characteristics (n = 51)

Characteristic	Mean ± SD
Gender (male / female)	23 / 28
Age (years)	40.27 ± 13.28
Height (cm)	167.54 ± 13.79
Weight (kg)	62.57 ± 11.45
LBP Duration (day)	51.78 ± 60.02
Oswestry Disability Index (score)	25.34 ± 9.61

agreement validity, the PIT and SIT were compared to the RF results. Spearman correlation analysis was calculated to provide validity on K-ODI, PIT, SIT, and RF. MedCalc statistical software was used for data analysis, considering p-values <.05 as statistically significant.

Ⅲ. Results

Among the 51 patients recruited, 2 denied RF evaluation and were excluded. There were 23 males and 28 females, with mean age, height, and weight of 40.27 13.28 years, 167.54 13.79 cm, and 62.57 11.45 kg, respectively. The mean LBP duration was 51.78 60.02 days, and mean K-ODI was 25.34 9.61 scores (Table 1). A STARD flowchart is provided in Fig. 1. No adverse events were reported.

The intra-rater reliability test shows the substantial value of sensitivity (PIT = .95, SIT = .98), specificity (PIT = .88, SIT = .70), kappa (PIT = .79, SIT = .73), and PABAK (PIT = .88, SIT = .84). The inter-rater reliability test appears to have a lower and reasonable value of sensitivity (PIT = .93, SIT = .98), specificity (PIT = .25, SIT = .80), kappa (PIT = .21, SIT = .80), and PABAK (PIT = .65, SIT = .88) (Table 2).

The validity between the PIT and SIT as compared to the RF shows sensitivity (PIT = .65, SIT = .68), specificity (PIT = .63, SIT = .70), kappa (PIT = .17, SIT = .28), PABAK (PIT = .29, SIT = .37) (Table 3).

The correlation coefficient among the PIT, SIT, RF,

Table 2. Intra-reliability and Inter-reliability of PIT and SIT (n = 51)

	Sensitivity	Specificity	PPV	NPV	Prevalence index	% agreement	Cohen's Kappa (CI)	PABAK
Intra-rater								
PIT	.95	.88	.98	.78	.84	94%	.79 (.56, 1.02)	.88
SIT	.98	.70	.93	.80	.80	92%	.73 (.48, .98)	.84
Inter-rater								
PIT	.93	.25	.87	.40	.84	82%	.21 (-.14, .56)	.65
SIT	.98	.80	.95	.88	.80	94%	.80 (.59, 1.02)	.88

PPV: positive predictive value, NPV: negative predictive value, PABAK: prevalence-adjusted bias-adjusted Kappa, PIT: prone instability test, SIT: side-lying instability test

Table 3. The Validity of PIT and SIT were Compared RF (n = 51)

	Sensitivity	Specificity	PPV	NPV	Prevalence index	% agreement	Cohen's Kappa (CI)	PABAK
PIT	.65	.63	.90	.25	.84	65%	.17 (-.07, .41)	.29
SIT	.68	.70	.90	.35	.80	68%	.28 (.03, .53)	.37

PPV: positive predictive value, NPV: negative predictive value, PABAK: prevalence-adjusted bias-adjusted Kappa, PIT: prone instability test, SIT: side-lying instability test

and K-ODI was r = .52~.72. The PIT was significantly correlated with RF (r = .69) and K-ODI (r = .53) while SIT with RF (r = .73) (Table 4).

IV. Discussion

We investigated the inter-rater and intra-rater reliability tests of the PIT and SIT. Participants who experienced pain during the PIT or SIT on the lumbar vertebra were classified as positive. We compared the K-ODI with the RF for validity and found that the PIT and SIT have high sensitivity and specificity in the intra-rater reliability test but high sensitivity and low specificity in the inter-rater reliability test. The PIT and SIT validities include good sensitivity and specificity compared to the RF. We confirmed the significant correlation among PIT, SIT, RF, and K-ODI (r = .52~.72).

In a previous study [10], the LBP subjects underwent PIT so that physical therapists can check their lumbar instability at prone position, which showed a 91%

Table 4. The Correlation with all Variables (n = 51)

	RF	K-ODI
PIT	.690*	.528*
SIT	.727*	.230

* correlation is significant at the .05 using spearman's rho
K-ODI: The score of Korean version Oswestry disability index,
PIT: The number of positive sign in prone instability test,
SIT: The number of positive sign in side-lying instability test,
RF: The number of positive sign in radiological finding

agreement and .87 reliability. Ravenna et al. [8] examined 30 LBP patients in a local community with two orthopedic manipulative therapists, which were all surveyed for their history, previous LBP experience, physical activity level, and general health status. The first examiner performed the SIT, while the second examiner conducted the PIT. This research statistic exhibited 63% agreement, .10 kappa, and .27 PABAK values. PABAK was calculated to evaluate the effect on kappa. Low kappa values were caused in part by a skewed distribution of individuals responses [8].

In the previous study by Schneider et al. [20], the PIT have a kappa and PABAK of .54 and .58, respectively, with 79% agreement. In this study, the PIT showed substantial results in the inter-rater kappa (.65) and PABAK (.77).

Physical therapists can determine the extent of the ligament sensitivity between the lumbar segments by examining the spine stability. The SIT was influenced by the tenderness of a ligament. Therefore, how the ligament sensitivity will be located in the LSI can be identified better.

As a result of this study, the SIT and PIT obtained high reliability. In a previous study [19], the K-ODI cut-off score of patients with LBP was 22.07. The average K-ODI of the subjects in this study was 25.33 ± 9.61. The SIT showed high validity compared to RF, which is a standard measurement in this study.

Clinical examination and decisions for patients with LBP require physiological and anatomical knowledge. Mechanical methods still have difficulties in identifying all normal or abnormal artificial movements of the patients to get the best evidence. LSI is an excellent examination tool to identify if patients with LBP have structural instability or not. The PIT is a typical test that predicts LSI. However, the PIT method has difficulty in correctly distinguishing muscle tenderness due to the difficulty in the case of each patient palpation. The use of the SIT, a modified prone instability test, instead of PIT in this study because the prone position is difficult for patients with LBP.

However, this study has some limitations. Firstly, we did not use the radiographic image on the flexion-extension position in the radiographic analysis to identify quantity abnormalities of motion, but we used a natural standing position image. Secondly, our result did not include all ages and recruit many patients with LBP and thus, does not have enough data to generalize.

Ⅴ. Conclusion

This study provides the first evidence reporting the SIT's reliability and validity, which modified the PIT for detecting LSI. The PIT and SIT had high specificity and sensitivity, which are easier to track in patients with pain and more flexible inability to handle the problem. Lumbar stability examination on the prone position requires a lot of practice to find the palpation's exact location. Compared with PIT, the SIT provides more comfort in tracking the patients' pain complaint as it copes with posture flexibility.

References

[1] Borkan JM, Koes B, Reis S, et al. A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. Spine (Phila Pa 1976). 1998;23(18):1992-6.

[2] Jang SY, Kong, MH, Hymanson HJ, Jin TK, et al. Radiographic Parameters of Segmental Instability in Lumbar Spine Using Kinetic MRI. J Korean Neurosurg Soc. 2009;45(1):24-31.

[3] O'Sullivan PB, Phyty GD, Twomey LT, et al. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. Spine (Phila Pa 1976). 1997;22(24): 2959-67.

[4] Engel GL. The need for a new medical model: a challenge for biomedicine. Psychodyn Psychiatry. 2012;40(3): 377-96.

[5] Abbott JH, McCane B, Herbison P, et al. Lumbar segmental instability: a criterion-related validity study of manual therapy assessment. BMC Musculoskelet Disord. 2005; 6(56):1-10

[6] Behrsin JF, Andrews FJ. Lumbar segmental instability: Manual assessment findings supported by radiological measurement (A case study). Aust J Physiother. 1991; 37(3):171-3.

[7] Kirkaldy-Willis WH, Farfan HF. Instability of the lumbar spine. Clin Orthop Relat Res. 1982;(165):110-23.

[8] Ravenna MM, Hoffman SL, Van Dillen LR. Low interrater reliability of examiners performing the prone instability test: a clinical test for lumbar shear instability. Arch Phys Med Rehabil. 2011;92(6):913-9.

[9] Ferrari S, Manni T, Bonetti F, et al. A literature review of clinical tests for lumbar instability in low back pain: validity and applicability in clinical practice. Chiropr Man Therap. 2015;23(14):1-12.

[10] Hicks GE, Fritz JM, Delitto A, et al. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. Arch Phys Med Rehabil. 2003;84(12):1858-64.

[11] Tonosu J, Takeshita K, Hara N, et al. The normative score and the cut-off value of the Oswestry Disability Index (ODI). Eur Spine J. 2012;21(8):1596-602.

[12] Robinson M. Clinical diagnosis and treatment of a patient with low back pain using the patient response model: A case report. Physiother Theory Pract. 2016;32(4): 315-23.

[13] Hasegawa K, Shimoda H, Kitahara K, et al. What are the reliable radiological indicators of lumbar segmental instability? J Bone Joint Surg Br. 2011;93(5):650-7.

[14] Alyazedi FM, Lohman EB, Wesley Swen R, et al. The

inter-rater reliability of clinical tests that best predict the subclassification of lumbar segmental instability: structural, functional and combined instability. J Man Manip Ther. 2015;23(4):197-204.

[15] Jeon CH, Kim DJ, Kim SK, et al. Validation in the cross-cultural adaptation of the Korean version of the Oswestry Disability Index. J Korean Med Sci. 2006;21(6): 1092-7.

[16] Alqarni AM, Schneiders AG, Hendrik PA. Clinical Tests to Diagnose Lumbar Segmental Instability: A Systematic Review. Journal of Orthopaedic & Sports Physical Therapy. 2011;41(3):130-40.

[17] Gopinath P. Lumbar segmental instability: Points to ponder. J Orthop. 2015;12(4):165-7.

[18] Ylinen J, Nykanen M, Kautiainen H, Kautiainen H, et al. Evaluation of repeatability of pressure algometry on the neck muscles for clinical use. Man Ther. 2007;12(2): 192-7.

[19] Hoehler FK, Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. J Clin Epidemiol. 2000;53(5):499-503.

[20] Schneider M, Erhard R, Brach J, et al. Spinal palpation for lumbar segmental mobility and pain provocation: an interexaminer reliability study. J Manipulative Physiol Ther. 2008;31(6):465-73.