Cureus
Part of SPRINGER NATURE

Open Access Original Article

Associations of Weight-Bearing Lunge Test Performance and Occiput-to-Wall Distance With Chronic Low Back Pain Intensity in Older Adults

Review began 10/11/2026
Review ended 10/21/2026
Published 10/27/2026

© **Copyright** 2026
Kuwabara et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.118700

Akira Kuwabara [1, 2], Kazuhide Inage [1], Masaomi Yamashita [3], Kohei Okuyama [3], Yuhei Iwatsuki [2], Yawara Eguchi [1, 4], Yasuhiro Shiga [1], Masahiro Inoue [1], Miyako Suzuki-Narita [5], Takahisa Hishiya [1], Masaya Mizutani [1], Susumu Tashiro [1], Shuhei Ohyama [1], Noritaka Suzuki [1], Kosuke Takeda [1], Akihiro Iida [1], Yu Otake [1], Sumihisa Orita [1, 6], Seiji Ohtori [1]

1. Department of Orthopedic Surgery, Graduate School of Medicine, Chiba University, Chiba, JPN 2. Department of Rehabilitation, Japan Community Health Care Organization (JCHO) Funabashi Central Hospital, Funabashi, JPN 3. Department of Orthopedic Surgery, Japan Community Health Care Organization (JCHO) Funabashi Central Hospital, Funabashi, JPN 4. Department of Orthopedic Surgery, Japan Community Health Care Organization (JCHO) Tokyo Joto Hospital, Tokyo, JPN 5. Department of Bioenvironmental Medicine, Graduate School of Medicine, Chiba University, Chiba, JPN 6. Center for Frontier Medical Engineering, Chiba University, Chiba, JPN

Corresponding author: Akira Kuwabara, debitrosi@gmail.com

Abstract

Purpose: We investigated whether weight-bearing lunge test (WBLT) performance and occiput-to-wall distance (OWD) were associated with chronic low back pain (LBP) intensity in older adults.

Materials and methods: This multicenter cross-sectional study included 208 outpatients with chronic LBP (mean age, 74.5 years) assessed between April 2024 and March 2025. LBP intensity was assessed using the Numerical Rating Scale (NRS). Multivariable linear regression analyses were performed with NRS score as the dependent variable. Model 1 included WBLT performance and OWD, whereas Model 2 additionally included age, sex, body mass index (BMI), symptom duration, and appendicular skeletal muscle mass (ASM) as potential confounders. Receiver operating characteristic (ROC) analyses for moderate-to-severe pain (NRS ⩾ 4) were conducted as exploratory secondary analyses.

Results: In Model 1, lower WBLT performance (standardized $\beta = -0.414$, $t = -5.470$, $p < 0.001$) and greater OWD (standardized $\beta = 0.190$, $t = 2.513$, $p = 0.013$) were significantly associated with greater LBP intensity. After additional adjustment in Model 2, lower WBLT performance remained significantly associated with greater LBP intensity (standardized $\beta = -0.465$, $t = -5.981$, $p < 0.001$), whereas the association between OWD and pain intensity was attenuated and was no longer statistically significant (standardized $\beta = 0.143$, $t = 1.720$, $p = 0.088$). Exploratory ROC analyses identified cutoff values of 0.9 cm for WBLT and 4.5 cm for OWD for moderate-to-severe pain.

Conclusion: In older adults with chronic LBP, lower WBLT performance remained significantly associated with greater pain intensity in the fully adjusted model, whereas the association between OWD and pain intensity was not statistically significant. However, OWD showed higher discriminatory ability in the exploratory ROC analysis and was significantly associated with pain intensity in an additional sensitivity analysis. Both measures may provide complementary clinical information, although the cross-sectional design and missing data warrant cautious interpretation.

Categories: Physical Medicine & Rehabilitation, Geriatrics, Orthopedics
Keywords: chronic low back pain, cross-sectional study, occiput-to-wall distance, older adults, pain intensity, weight-bearing lunge test

Introduction

Low back pain (LBP) is a common musculoskeletal condition, with a reported lifetime prevalence of up to 84% and an estimated chronicity rate of approximately 23% [1]. When accompanied by activity limitations, LBP can substantially affect quality of life and impose considerable healthcare and socioeconomic burdens, and its multifactorial nature supports clinical assessment beyond the lumbar region alone [1].

Interactions between the lumbar spine and adjacent regions of the lower extremities have increasingly been recognized. Associations between hip disorders and lumbar symptoms have been described as hip-spine syndrome, whereas coexistence of knee and spinal disorders has been discussed in the context of knee-spine syndrome [2,3]. The possible contribution of ankle function to LBP, particularly in older adults, remains insufficiently characterized [4,5].

Limited ankle dorsiflexion during weight-bearing tasks may modify movement throughout the lower-

extremity kinetic chain. Previous studies have described changes in knee, hip, pelvic, and trunk movement associated with restricted dorsiflexion, suggesting a potential influence on mechanical demands placed on the lumbopelvic region [4-7]. The weight-bearing lunge test (WBLT) provides a simple clinical method for evaluating weight-bearing movement capacity [8]. In the present study, WBLT performance was interpreted as a weight-bearing functional measure rather than as an isolated measure of ankle dorsiflexion.

Postural characteristics were also considered in the present study. The occiput-to-wall distance (OWD) is an easily obtained clinical measure that has been used to identify thoracic hyperkyphotic posture, although it should not be regarded as a direct measure of global spinal alignment because the measured distance may also reflect cervical and lumbar alignment, pelvic compensation, and structural spinal changes [9].

Although previous research has examined relationships between lower-extremity alignment or mobility and LBP [4,5], to our knowledge, the associations of WBLT performance and OWD with chronic LBP intensity in older adults have not been well characterized. Accordingly, the present study examined whether WBLT performance and OWD were associated with LBP intensity in older adults. We hypothesized that poorer WBLT performance and greater OWD would be associated with greater pain intensity.

Materials And Methods

Study design

This multicenter cross-sectional study was conducted at seven facilities in Japan (Japan Community Health Care Organization (JCHO) Funabashi Central Hospital, Chiba University Hospital, Nishikawa Orthopedics Clinic, Kitachiba Spine & Sports Clinic (Inage Clinic), Sanmu Medical Center, JCHO Tokyo Joto Hospital, and Oyumino Central Hospital). The study protocol was approved by the Institutional Review Board of the School of Medicine, Chiba University (approval no. 2805) and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants before study participation.

Participants

Consecutive outpatients with chronic LBP who attended the participating facilities between April 2024 and March 2025 were recruited. Eligibility was assessed using the predefined inclusion and exclusion criteria described below.

Inclusion criteria

Participants were eligible for inclusion if they were aged 65 years or older, had a primary complaint of LBP persisting for at least three months, were able to ambulate independently and perform the WBLT, and were able to provide written informed consent for study participation.

Exclusion criteria

Participants were excluded if they had spinal disorders accompanied by neurological abnormalities, including radiculopathy, motor weakness, sensory deficits, diminished tendon reflexes, or imaging-confirmed nerve compression; were unable to cooperate with the assessments because of severe cognitive impairment or a similar condition; had undergone spinal surgery within the previous six months; or had ankle disorders that prevented appropriate performance of the WBLT, including post-traumatic deformity, severe osteoarthritis, or marked restriction of range of motion. The presence of spinal disorders accompanied by neurological abnormalities, including radiculopathy, was determined primarily based on physician diagnoses, medical records, and a review of radiographic and magnetic resonance imaging findings for all participants. These assessments were used to identify clinically recognized neurological abnormalities and nerve compression.

Participants were also excluded if they had a history of osteoporotic vertebral fractures or knee or hip conditions substantially affecting pain or mobility, including severe knee or hip pain requiring treatment within the preceding six months, a history of knee or hip joint replacement, or marked limitations in daily walking. Medical records, clinical history, and radiographic or magnetic resonance imaging findings were reviewed for all participants to identify osteoporotic vertebral fractures.

Mild ankle osteoarthritis or ankle pain was not considered an exclusion criterion if the participant was able to perform the WBLT. Eligibility was determined according to a standardized protocol shared across all participating facilities. A flowchart summarizing the participant recruitment and selection process is presented in Figure *1*.

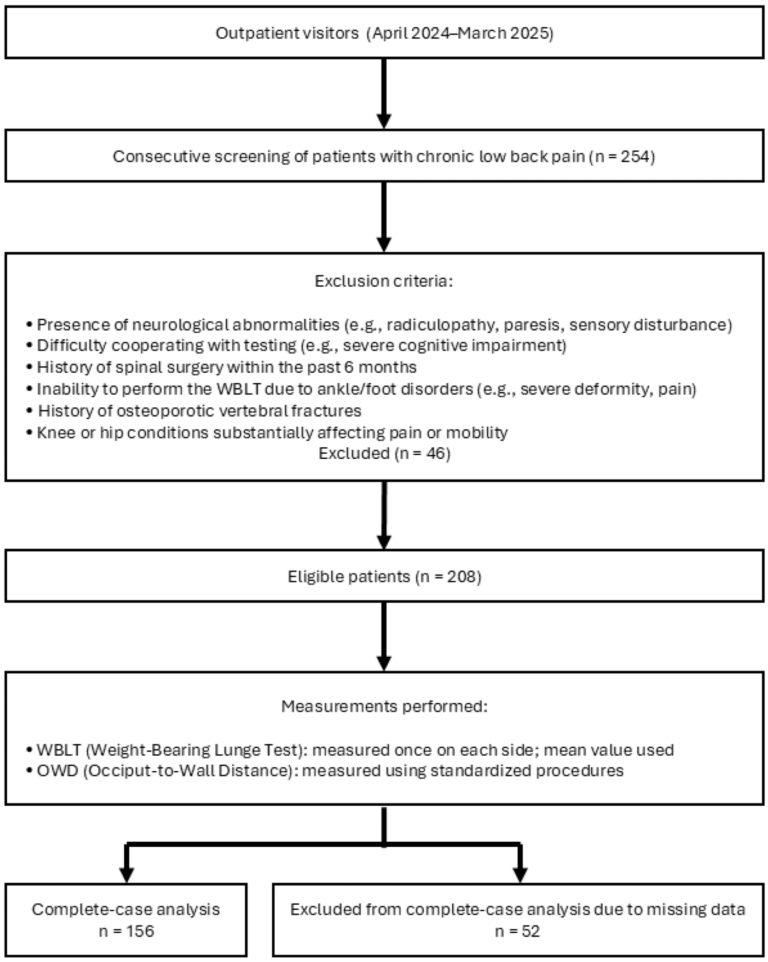


FIGURE 1: Flowchart of participant recruitment based on the predefined inclusion and exclusion criteria.

Sample size calculation

The required sample size for the primary analysis was estimated a priori using G*Power 3.1 (Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany) for multivariable linear regression with NRS as the dependent variable and WBLT performance as the primary explanatory variable [10]. For the sample size calculation, an effect size of $f^2 = 0.10$, a two-sided α level of 0.05, and a statistical power of 0.80 were assumed. The calculated minimum required sample size was approximately 151 participants.

Participating facilities and sample distribution

A total of 208 participants were enrolled across seven facilities. The number of participants at each facility was as follows: Facility A, 40; Facility B, 35; Facility C, 30; Facility D, 28; Facility E, 25; Facility F, 25; and Facility G, 25.

Outcome measures

The primary explanatory variables were WBLT performance and OWD, and the dependent variable was LBP intensity assessed using the NRS. Covariates included age, sex, body mass index (BMI), symptom duration, and appendicular skeletal muscle mass (ASM).

These covariates were selected a priori based on previous literature and clinical relevance. ASM was assessed using bioelectrical impedance analysis (BIA) and was included as a covariate because of its potential association with physical function and LBP in older adults [11]. LBP intensity was assessed using an 11-point Numerical Rating Scale (NRS), ranging from 0 (no pain) to 10 (worst imaginable pain). Participants reported their LBP intensity on the day of their visit to the participating facility.

Assessment procedures and bias mitigation

At each participating facility, assessments were performed by one or two physical therapists who received standardized training based on a unified protocol. The training consisted of a 30-minute didactic session covering the WBLT and OWD measurement procedures, practical training using simulated participants with five practice trials for each assessment, and confirmation of standardized measurement and recording procedures.

Weight-bearing lunge test

Participants stood barefoot facing a wall with the tested foot positioned on a measuring tape. The knee was advanced toward the wall in line with the third toe while the heel remained in contact with the floor. The maximum distance from the tip of the great toe to the wall at which the knee could touch the wall without lifting the heel was recorded to the nearest 0.1 cm.

Measurements were obtained once on each side, and the mean of the bilateral measurements was used for analysis. Because the WBLT is performed in a closed kinetic chain, the measured distance was interpreted in the present study as a weight-bearing functional measure rather than as an isolated measure of ankle dorsiflexion (Figure *2*).

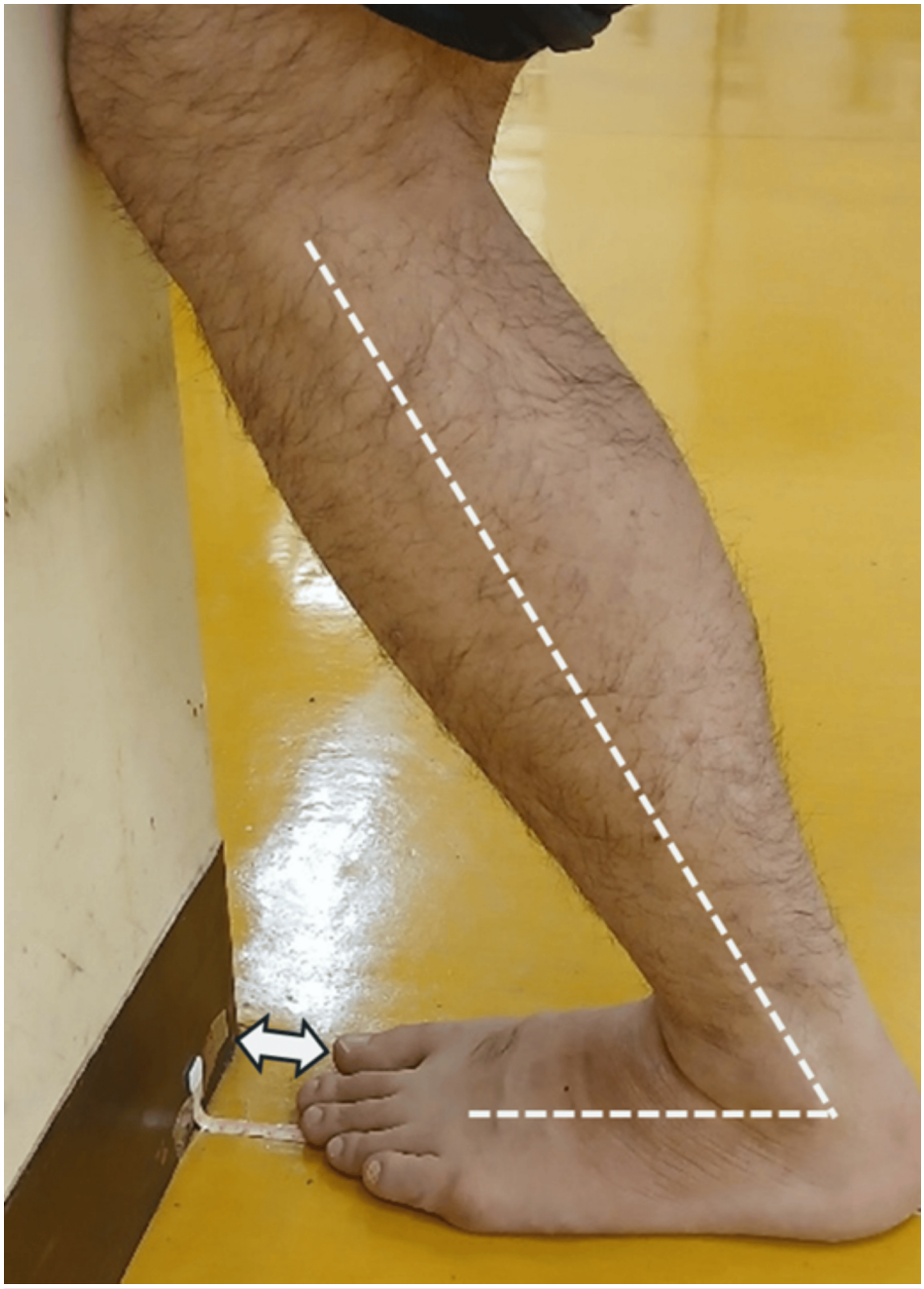


FIGURE 2: Assessment of weight-bearing lunge test (WBLT) performance.

Participants lunged toward a wall while keeping the heel on the floor, and the maximum distance from the tip of the great toe to the wall was measured.

Occiput-to-wall distance

Participants stood with their heels, sacrum, and thoracic spine in contact with the wall while maintaining a neutral head position. The distance between the external occipital protuberance and the wall was measured using two rigid rulers (Figure 3) [9].

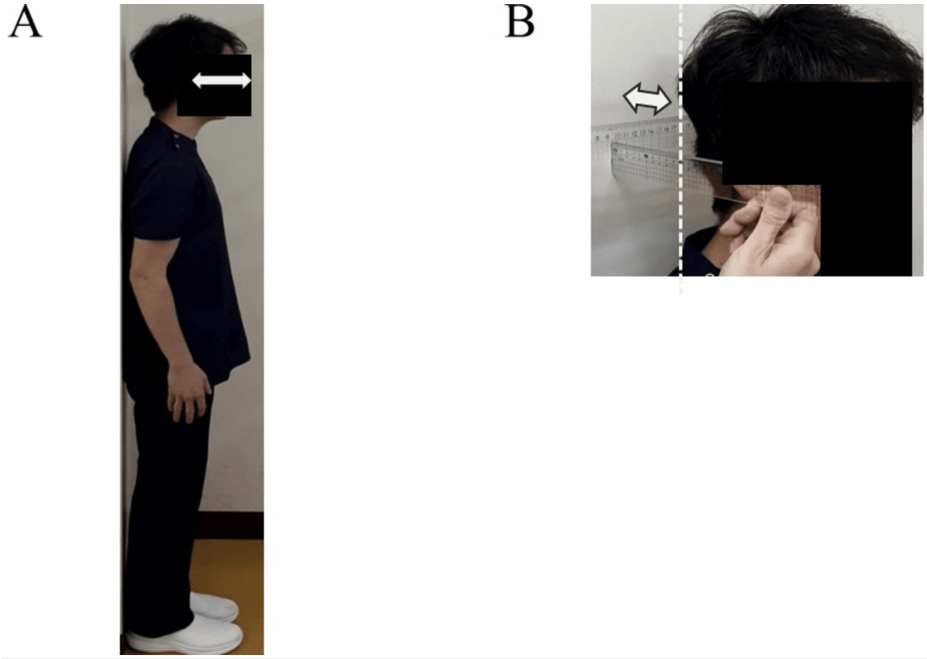


FIGURE 3: Assessment using the occiput-to-wall distance (OWD).

A: standing position during OWD measurement; B: actual measurement of the OWD using two rigid rulers. Participants stood upright with their heels, sacrum, and thoracic spine against the wall while maintaining a neutral head position. The distance between the external occipital protuberance and the wall was measured.

OWD was interpreted in the present study as a simple clinical measure of postural characteristics rather than as a direct measure of global spinal alignment. Measurements were performed using a standardized protocol, and, when feasible, the same assessor within each facility performed the measurements to reduce interrater variability.

Statistical analysis

Statistical analyses were performed using R software (R Foundation for Statistical Computing, Vienna, Austria). The normality of continuous variables was assessed using the Shapiro-Wilk test. For secondary comparisons between the mild pain group (NRS < 4) and the moderate-to-severe pain group (NRS ⩾ 4) [12,13], continuous variables were compared using the Wilcoxon rank-sum (Mann-Whitney) test, whereas categorical variables were compared using the chi-square test.

For the primary analysis, multivariable linear regression analyses were performed with NRS score as the dependent variable. Model 1 included WBLT performance and OWD as explanatory variables. Model 2 additionally included age, sex, BMI, symptom duration, and ASM as potential confounders. Regression analyses were performed using complete cases with available data for all variables included in each model. As a sensitivity analysis for missing data, multiple imputation by chained equations was performed using 40 imputed datasets in Python (Python Software Foundation, Fredericksburg, VA) (scikit-learn). Missing values were imputed using iterative Bayesian ridge regression with posterior sampling and a maximum of 20 iterations per imputation. The imputation model included NRS, WBLT, OWD, age, sex, BMI, symptom duration, and ASM. Symptom duration was converted from days to months by dividing by 30. The fully adjusted linear regression model was fitted to each imputed dataset, and the estimates were pooled using Rubin's rules.

Multicollinearity was assessed using variance inflation factors (VIFs). Additional sensitivity analyses were conducted to evaluate the robustness of the regression findings. The distribution of NRS scores and the residuals of the fully adjusted linear regression model were examined. Residual normality was assessed using the Shapiro-Wilk test, and heteroskedasticity was assessed using the Breusch-Pagan test. To account for heteroskedasticity, the fully adjusted linear regression model was re-evaluated using HC3 heteroskedasticity-consistent standard errors.

To examine the potential influence of differences across the seven participating facilities, sensitivity analyses were performed using cluster-robust standard errors at the facility level and a linear regression model including facility fixed effects. Facility membership was reconstructed from the original facility-grouped order of the dataset and the recorded number of participants at each facility. These additional analyses were conducted using Python (statsmodels).

Receiver operating characteristic (ROC) analyses were conducted as exploratory secondary analyses to evaluate the ability of WBLT performance and OWD to discriminate moderate-to-severe pain (NRS ⩾ 4). Areas under the curve (AUCs) and their 95% confidence intervals were calculated using the DeLong method. Optimal cutoff values were determined using the Youden index. Internal validation of the ROC-derived cutoffs was performed using 2,000 bootstrap resamples. In each bootstrap sample, the cutoff maximizing the Youden index was selected, and its performance was evaluated in both the bootstrap sample and the original dataset. Optimism was estimated as the mean difference in the Youden index between these evaluations and was subtracted from the apparent Youden index to obtain an optimism-corrected estimate. These analyses were performed using Python (scikit-learn).

Because lower WBLT values were associated with greater pain severity, the direction of the ROC analysis was specified accordingly. All statistical tests were two-sided, and a p-value < 0.05 was considered statistically significant.

Results

Participant characteristics

A total of 208 participants were included in the overall study cohort, including 136 men (65.4%) and 72 women (34.6%). The mean age of the participants was 74.5 ± 6.7 years. For secondary between-group comparisons, participants were classified into a mild pain group (NRS < 4, n = 88; age, 75.0 ± 6.6 years) and a moderate-to-severe pain group (NRS ⩾ 4, n = 120; age, 74.1 ± 6.8 years). No significant between-group differences were observed in age, sex, BMI, symptom duration, or ASM (Table *1*).

Variable	All cases (n = 208)	Mild pain (NRS < 4, n = 88)	Moderate-to-severe pain (NRS ≥ 4, n = 120)	Test statistic	p-value
Age (years), mean ± SD	74.5 ± 6.7	75.0 ± 6.6	74.1 ± 6.8	W = 5568	0.501
Sex, n (%)	-	-	-	χ^2 = 0.527	0.468
Male	136 (65.4)	60 (68.2)	76 (63.3)	-	-
Female	72 (34.6)	28 (31.8)	44 (36.7)	-	-
BMI (kg/m²), mean ± SD	24.0±4.2	24.4±3.6	23.6±4.5	W = 5736	0.130
Duration of symptoms (months), mean ± SD	29.6±21.3	26.4±19.5	31.8±22.2	W = 4336	0.086
NRS, mean ± SD	4.5 ± 3.4	1.0 ± 1.3	7.1 ± 1.9	W = 0	<0.001
WBLT (cm), mean ± SD	2.9±3.4	4.8±3.7	1.5±2.3	W = 7678	<0.001
OWD (cm), mean ± SD	6.0±5.9	2.6±2.3	8.4±6.4	W = 1672	<0.001
ASM (kg), mean ± SD	43.1±9.3	43.7±8.5	42.8±9.7	W = 3744	0.175

TABLE 1: Patient characteristics stratified by pain severity.

Data are presented as mean ± standard deviation (SD) for continuous variables and n (%) for categorical variables. Continuous variables were compared using the Wilcoxon rank-sum (Mann–Whitney) test, and categorical variables were compared using the Pearson chi-square test. Test statistics are reported as W for the Wilcoxon rank-sum test and χ^2 for the chi-square test. All statistical tests were two-sided, and p < 0.05 was considered statistically significant. A hyphen (-) indicates that the value is not applicable.

ASM: appendicular skeletal muscle mass; BIA: bioelectrical impedance analysis; BMI: body mass index; NRS: Numerical Rating Scale; OWD: occiput-to-wall distance; WBLT: weight-bearing lunge test

The number of observations included in each comparison varied because of missing data. The Wilcoxon rank-sum tests were performed using all available observations for each variable.

Physical function measures

WBLT and OWD results are presented in Table *1*. WBLT performance was significantly lower in the moderate-to-severe pain group than in the mild pain group (1.5 ± 2.3 cm vs. 4.8 ± 3.7 cm, W = 7678, p < 0.001). OWD was significantly greater in the moderate-to-severe pain group than in the mild pain group (8.4 ± 6.4 cm vs. 2.6 ± 2.3 cm, W = 1672, p < 0.001). The median WBLT values were 5.0 cm (interquartile range (IQR), 1.125-8.6 cm) in the mild pain group and 0.0 cm (IQR, 0.0-3.5 cm) in the moderate-to-severe pain group. WBLT values of 0 cm were observed in 22 of 88 participants (25.0%) in the mild pain group and 74 of

116 participants (63.8%) with available WBLT data in the moderate-to-severe pain group, suggesting a potential floor effect. A histogram of WBLT performance is presented in Figure 4. Of the 204 participants with available WBLT measurements, 96 (47.1%) had a value of 0 cm, indicating a substantial concentration of observations at the lower end of the measurement range.

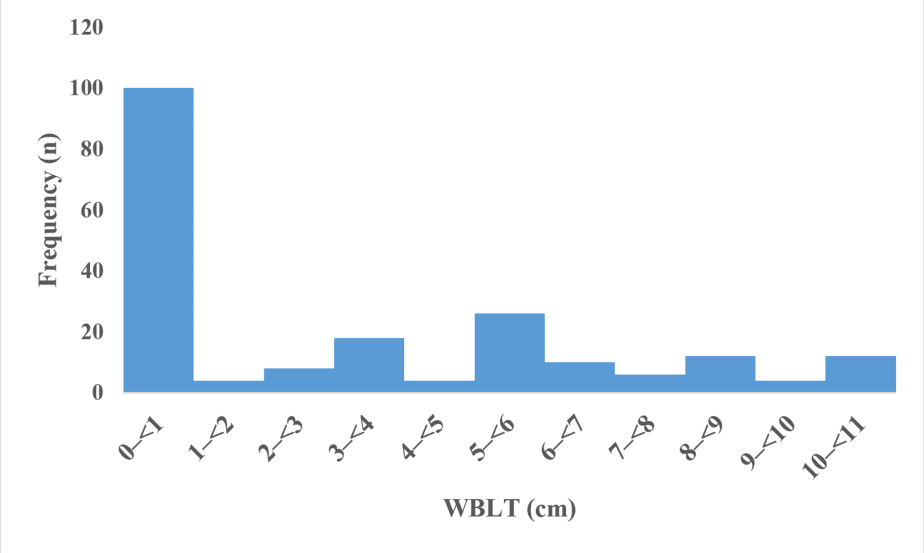


FIGURE 4: Distribution of weight-bearing lunge test (WBLT) performance.

The histogram illustrates the distribution of WBLT measurements among 204 participants with available data, using 1-cm intervals. Of these, 96 participants (47.1%) had a WBLT value of 0 cm, indicating a substantial concentration of measurements at the lower end of the range.

Multivariable linear regression analysis

Complete-case regression analyses included 156 participants. In Model 1, which included WBLT performance and OWD, lower WBLT performance was significantly associated with greater LBP intensity (standardized $\beta = -0.414$, $t = -5.470$, $p < 0.001$), and greater OWD was also significantly associated with greater LBP intensity (standardized $\beta = 0.190$, $t = 2.513$, $p = 0.013$). The model explained 27.2% of the variance in NRS scores ($R^2 = 0.272$, adjusted $R^2 = 0.262$; $F(2,153) = 28.56$, $p < 0.001$).

In Model 2, after adjustment for age, sex, BMI, symptom duration, and ASM, lower WBLT performance remained significantly associated with greater LBP intensity (standardized $\beta = -0.465$, $t = -5.981$, $p < 0.001$), whereas the association between OWD and LBP intensity was no longer statistically significant (standardized $\beta = 0.143$, $t = 1.720$, $p = 0.088$). The adjusted model explained 31.1% of the variance in NRS scores ($R^2 = 0.311$, adjusted $R^2 = 0.279$; $F(7,148) = 9.55$, $p < 0.001$). The regression results are summarized in Table 2. The VIFs ranged from 1.146 to 4.327, with all values below 5, suggesting no substantial multicollinearity according to this criterion.

Variables	Standardized β	B (95% CI)	t statistic	p-value
Model 1				
WBLT (cm)	−0.414	−0.404 (−0.549 to −0.259)	−5.470	<0.001
OWD (cm)	0.190	0.098 (0.021 to 0.175)	2.513	0.013
Model 2				
WBLT (cm)	−0.465	−0.454 (−0.604 to −0.304)	−5.981	<0.001
OWD (cm)	0.143	0.074 (−0.011 to 0.159)	1.720	0.088
Age (years)	−0.121	−0.061 (−0.134 to 0.012)	−1.651	0.100
Sex (male)	−0.185	−1.308 (−3.121 to 0.505)	−1.426	0.156
BMI (kg/m²)	−0.029	−0.022 (−0.148 to 0.104)	−0.345	0.731
Symptom duration	−0.118	−0.018 (−0.041 to 0.005)	−1.547	0.124
ASM (kg)	0.128	0.047 (−0.055 to 0.149)	0.911	0.364

TABLE 2: Multivariable linear regression analyses of the associations of WBLT performance and OWD with pain intensity (NRS).

Dependent variable: NRS score. Model 1 included WBLT performance and OWD. Model 2 additionally included age, sex, BMI, symptom duration, and ASM as covariates. B represents the unstandardized regression coefficient, and β represents the standardized regression coefficient. Test statistics for individual regression coefficients are reported as t values. The overall model test was evaluated using the F statistic. All statistical tests were two-sided, and $p < 0.05$ was considered statistically significant.

ASM: appendicular skeletal muscle mass; BMI: body mass index; NRS: Numerical Rating Scale; OWD: occiput-to-wall distance; WBLT: weight-bearing lunge test

In sensitivity analyses using the same 156 complete cases and covariates as Model 2, the association between WBLT performance and NRS scores remained statistically significant when WBLT was log-transformed using log(1 + WBLT) ($p < 0.001$) or categorized as 0 cm versus greater than 0 cm ($p < 0.001$). These findings suggest that the association was not limited to the original continuous specification of WBLT. Of the 208 participants, 156 (75.0%) had complete data for the multivariable regression analyses, whereas 52 (25.0%) had missing data in at least one required variable. Missing values were observed for BMI (n = 4), symptom duration (n = 4), WBLT (n = 4), OWD (n = 8), and ASM (n = 40), with some overlap between variables. Missing ASM data (n = 40) resulted from BIA measurements not being performed due to participant-related circumstances. Participants included in the complete-case analysis had higher NRS scores than those excluded (4.97 ± 3.35 vs. 3.08 ± 3.35).

In an additional sensitivity analysis excluding ASM from the covariates, 196 participants were included. WBLT performance remained significantly associated with NRS scores (B = −0.350, 95% CI: −0.497 to −0.203, $p < 0.001$). OWD was also significantly associated with NRS scores in this model (B = 0.125, 95% CI: 0.041 to 0.209, $p = 0.004$). The distribution of NRS scores showed a concentration of observations at 0 and at higher pain scores. Diagnostic tests of the fully adjusted linear regression model indicated departures from residual normality (Shapiro-Wilk test, $p < 0.001$) and homoscedasticity (Breusch-Pagan test, $p < 0.001$). Nevertheless, WBLT performance remained significantly associated with NRS scores when HC3 heteroskedasticity-consistent standard errors were used ($p < 0.001$).

In sensitivity analyses accounting for participating facilities, the association between WBLT performance and NRS scores remained statistically significant when cluster-robust standard errors were applied at the facility level (B = −0.454, $p = 0.021$). The association also remained significant in a model including facility fixed effects (B = −0.554, $p < 0.001$). OWD was significantly associated with NRS scores in the facility fixed-effects model (B = 0.114, $p = 0.007$). In the multiple imputation sensitivity analysis including all 208 participants, WBLT performance remained significantly associated with NRS scores (B = −0.362, 95% CI: −0.507 to −0.217, $p < 0.001$). OWD was also significantly associated with NRS scores (B = 0.127, 95% CI: 0.044 to 0.211, $p = 0.003$). These findings suggest that the association between WBLT performance and pain intensity persisted after multiple imputation, whereas the association between OWD and pain intensity differed between the complete-case and multiple imputation analyses.

ROC analysis

Exploratory ROC analyses were performed to evaluate the ability of WBLT performance and OWD to discriminate moderate-to-severe pain (NRS ⩾ 4). The ROC analyses included 204 participants for WBLT and 200 participants for OWD.

For WBLT, the optimal cutoff value determined using the Youden index was 0.9 cm, with an AUC of 0.75 (95% CI, 0.69-0.82), sensitivity of 67.2%, and specificity of 75.0%. For OWD, the optimal cutoff value was 4.5 cm, with an AUC of 0.83 (95% CI, 0.77-0.88), sensitivity of 82.8%, and specificity of 71.4%. The ROC analysis results are summarized in Table 3. In the bootstrap internal validation with 2,000 resamples, the estimated optimism in the Youden index was 0.032 for both WBLT and OWD. The corresponding optimism-corrected Youden indices were 0.390 and 0.510, respectively. Bootstrap resampling also indicated variability in the selected cutoff values, particularly for WBLT, suggesting that the derived thresholds should be interpreted cautiously.

Variable	n	Cutoff value (cm)	AUC (95% CI)	Sensitivity (%)	Specificity (%)
WBLT (cm)	204	0.9	0.75 (0.69–0.82)	67.2	75.0
OWD (cm)	200	4.5	0.83 (0.77–0.88)	82.8	71.4

TABLE 3: Exploratory ROC analyses for moderate-to-severe pain (NRS ≥ 4).

Data are presented as cutoff values, area under the curve (AUC) with 95% confidence intervals (CI), sensitivity, and specificity. ROC analyses were performed to evaluate the ability of WBLT performance and OWD to discriminate moderate-to-severe pain, defined as NRS ≥ 4. AUCs and 95% CIs were calculated using the DeLong method, and optimal cutoff values were determined using the Youden index.

AUC: area under the curve; CI: confidence interval; NRS: Numerical Rating Scale; OWD: occiput-to-wall distance; ROC: receiver operating characteristic; WBLT: weight-bearing lunge test

Discussion

In this multicenter cross-sectional study, we examined the associations of WBLT performance and OWD with LBP intensity in older adults with chronic LBP. In Model 1, lower WBLT performance and greater OWD were both significantly associated with higher NRS scores. After additional adjustment for age, sex, BMI, symptom duration, and ASM in Model 2, lower WBLT performance remained significantly associated with greater LBP intensity, whereas the association between OWD and LBP intensity was attenuated and was no longer statistically significant. These findings suggest that WBLT performance may provide clinically relevant information regarding pain intensity in older adults with chronic LBP, while the association between OWD and pain intensity may be more strongly influenced by other patient characteristics.

Lower WBLT performance remained significantly associated with greater LBP intensity after adjustment for age, sex, BMI, symptom duration, and ASM. In Model 2, a 1-cm lower WBLT value was associated with an approximately 0.45-point higher NRS score (B = −0.454). This estimate quantifies the observed relationship between WBLT performance and pain intensity in the present sample. However, this association should not be interpreted as evidence that isolated ankle restriction directly determines pain severity.

One possible explanation for the observed association is that reduced weight-bearing ankle motion may be accompanied by compensatory movement elsewhere in the lower-extremity kinetic chain. Previous studies have shown that limited ankle dorsiflexion can be associated with alterations in knee, hip, pelvic, and trunk movement during weight-bearing tasks, potentially modifying mechanical demands across the lumbopelvic region [4-6]. Nevertheless, because the test is performed in a closed kinetic chain, WBLT performance should not be interpreted as an isolated measure of ankle dorsiflexion. Accordingly, our results indicate an association between overall weight-bearing functional performance and LBP intensity rather than demonstrating a direct causal pathway between restricted ankle dorsiflexion and greater pain.

A previous study involving female students aged 18-30 years reported no significant association between limited ankle dorsiflexion and non-specific mechanical LBP [5]. This finding differs from the significant association observed in the present study of older adults with chronic LBP. Differences in age, participant characteristics, study design, and clinical setting may have contributed to the discrepant results. Because these studies examined distinct populations under different methodological conditions, the findings should not be interpreted as evidence that age itself explains the difference in associations. Nevertheless, our results suggest that assessment of weight-bearing lower-extremity function may provide clinically relevant information when evaluating older adults with chronic LBP.

Greater OWD was significantly associated with higher LBP intensity in Model 1, but this association was attenuated and was no longer statistically significant after adjustment for age, sex, BMI, symptom duration, and ASM in Model 2. This change suggests that the relationship between OWD and pain intensity may be

influenced by other patient characteristics included in the adjusted model. Although OWD is a simple clinical measure that has been used to assess thoracic hyperkyphotic posture, it should not be regarded as a direct measure of overall spinal alignment because its value may also reflect contributions from cervical and lumbar alignment, pelvic compensation, and structural spinal characteristics [9]. Therefore, in the present study, OWD may have represented a broader postural characteristic rather than an independent correlate of LBP intensity. Age-related kyphosis, including senile kyphosis, may contribute to increased OWD and has been associated with neck and back symptoms in older adults [14]. Therefore, age-related postural changes should be considered when interpreting the association between OWD and LBP intensity.

In the exploratory ROC analyses, both WBLT performance and OWD showed discriminatory ability for moderate-to-severe pain defined as NRS ⩾ 4. The optimal cutoff values identified in this study were 0.9 cm for WBLT and 4.5 cm for OWD. Because these values were derived from a single cross-sectional cohort, they should be considered exploratory estimates rather than established clinical thresholds. Although bootstrap internal validation was performed, variability in the selected cutoff values, particularly for WBLT, indicates that the stability and generalizability of these thresholds remain uncertain.

They should therefore not be used as diagnostic criteria or as direct targets for treatment. Validation in independent samples, preferably using prospective study designs, is necessary before their clinical applicability can be determined. Notably, OWD demonstrated a higher AUC than WBLT performance in the exploratory ROC analysis (0.83 vs. 0.75), despite its association with pain intensity not reaching statistical significance in the fully adjusted regression model. These findings are not necessarily contradictory because ROC analysis evaluates the ability of each measure to discriminate between pain severity groups, whereas multivariable linear regression examines associations with continuous pain intensity after adjustment for covariates. Furthermore, OWD was significantly associated with pain intensity in the additional sensitivity analysis excluding ASM, suggesting that its association may depend on the covariates included and the analyzed sample. Therefore, the present findings do not establish the superiority of WBLT over OWD, and both measures may provide complementary clinical information. Further studies are required to clarify their respective roles.

Several limitations should be considered when interpreting these findings. First, the cross-sectional design does not allow conclusions regarding temporal or causal relationships. Lower WBLT performance and greater OWD therefore cannot be regarded as causes of more intense LBP, and reverse causation remains possible. Second, LBP intensity was evaluated using the NRS, a self-reported measure that may be affected by individual perception and other subjective factors. Third, potentially relevant psychosocial variables, including fear-avoidance beliefs, depressive symptoms, and pain catastrophizing, were not assessed and may have contributed to residual confounding [15,16]. Fourth, WBLT and OWD are indirect clinical measures and should not be interpreted as isolated measures of ankle mobility or global spinal alignment, respectively. A potential floor effect in WBLT performance should also be considered, particularly among participants with greater pain intensity. A substantial proportion of participants had measurements at the lower end of the scale, which may have limited the ability of the WBLT to distinguish differences in weight-bearing functional performance among individuals with moderate-to-severe pain. This distributional characteristic should be considered when interpreting the observed associations.

In addition, WBLT measurements were obtained only once per side, and the mean of the bilateral measurements was used for analysis. This approach may have increased measurement variability and obscured clinically relevant side-to-side differences. Furthermore, although assessors received standardized training, inter-rater reliability across the participating facilities was not evaluated. These measurement-related limitations should be considered when interpreting the findings.

Fifth, although imaging findings were reviewed to identify osteoporotic vertebral fractures, quantitative radiographic measurements of spinal and pelvic alignment were not performed, limiting our ability to directly evaluate the proposed biomechanical relationships.

Although participants with severe knee or hip conditions affecting pain or mobility were excluded, individuals with mild or asymptomatic knee or hip osteoarthritis may have been included. These conditions may have influenced WBLT performance and contributed to residual confounding. In addition, although participants with a history of osteoporotic vertebral fractures were excluded based on clinical history, medical records, and imaging findings, the possibility of undetected vertebral fractures cannot be completely ruled out. Therefore, the potential influence of these musculoskeletal conditions should be considered when interpreting the associations of WBLT performance and OWD with LBP intensity. Furthermore, other potentially relevant factors, including quadriceps muscle weakness, fear of movement, spinal stenosis without neurological abnormalities, and analgesic use, were not included as covariates in the multivariable analysis and may have contributed to residual confounding. Although spinal disorders accompanied by neurological abnormalities were excluded based on physician diagnoses, medical records, and imaging findings, the possibility of unrecognized or atypical radiculopathy cannot be completely ruled out. Another important limitation is missing data. The primary multivariable analyses were restricted to 156 of 208 participants, and those included had higher pain intensity than those excluded. This difference raises concerns about potential selection bias. Missing ASM data were primarily attributable to participant-related circumstances that prevented BIA measurements. Because the missingness mechanism could not be

definitively established, the possibility of missing-not-at-random data cannot be excluded. Although multiple imputation was performed as a sensitivity analysis, its results should be interpreted in light of the underlying assumptions regarding missing data.

Although an additional analysis excluding ASM allowed the inclusion of 196 participants and the association between WBLT performance and pain intensity remained significant, this analysis did not account for muscle mass. Furthermore, the association between OWD and pain intensity differed between models. These findings highlight the need for cautious interpretation of the regression results. Although the association between WBLT performance and pain intensity remained statistically significant in additional sensitivity analyses, several statistical limitations should be acknowledged. The original linear regression model showed departures from residual normality and homoscedasticity. Although the association between WBLT performance and pain intensity remained statistically significant when HC3 heteroskedasticity-consistent standard errors were used, departures from the assumptions of the original linear regression model warrant cautious interpretation.

Furthermore, although facility-level sensitivity analyses were conducted, only seven facilities were included. Cluster-robust standard errors may be less reliable with a small number of clusters, and the estimated associations may still be affected by unmeasured facility-level differences. Therefore, these supplementary analyses should be interpreted cautiously.

Finally, the ROC-derived cutoff values were generated from the present cohort and require validation in independent populations before being considered for clinical use.

Overall, the present findings indicate that lower WBLT performance is associated with greater LBP intensity in older adults even after adjustment for several potential confounders, whereas the association between OWD and pain intensity is less robust after adjustment. These results suggest that weight-bearing lower-extremity function may provide clinically relevant information as part of the broader assessment of older adults with chronic LBP. However, because the study was cross-sectional, the direction and temporal sequence of these associations cannot be determined. Prospective studies are needed to clarify whether changes in WBLT performance are related to subsequent changes in pain intensity or other clinically important outcomes.

Conclusions

In older adults with chronic LBP, lower WBLT performance remained significantly associated with greater pain intensity after adjustment for age, sex, BMI, symptom duration, and ASM. Although the association between OWD and pain intensity was not statistically significant in the fully adjusted model, OWD demonstrated higher discriminatory ability in the exploratory ROC analysis and was significantly associated with pain intensity in an additional sensitivity analysis. These findings suggest that WBLT performance and OWD may provide complementary information when assessing older adults with chronic LBP. However, the cross-sectional design, potential floor effect in WBLT performance, and missing data limit the interpretation of these findings. Further prospective studies are needed to clarify their clinical significance.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Akira Kuwabara, Kazuhide Inage, Masaomi Yamashita, Yawara Eguchi, Yasuhiro Shiga, Masahiro Inoue, Sumihisa Orita, Seiji Ohtori

Acquisition, analysis, or interpretation of data: Akira Kuwabara, Kazuhide Inage, Masaomi Yamashita, Kohei Okuyama, Yuhei Iwatsuki, Yawara Eguchi, Yasuhiro Shiga, Masahiro Inoue, Miyako Suzuki-Narita, Takahisa Hishiya, Masaya Mizutani, Susumu Tashiro, Shuhei Ohyama, Noritaka Suzuki, Kosuke Takeda, Akihiro Iida, Yu Otake, Sumihisa Orita, Seiji Ohtori

Drafting of the manuscript: Akira Kuwabara, Kazuhide Inage

Supervision: Kazuhide Inage, Seiji Ohtori

Critical review of the manuscript for important intellectual content: Masaomi Yamashita, Kohei Okuyama, Yuhei Iwatsuki, Yawara Eguchi, Yasuhiro Shiga, Masahiro Inoue, Miyako Suzuki-Narita, Takahisa Hishiya, Masaya Mizutani, Susumu Tashiro, Shuhei Ohyama, Noritaka Suzuki, Kosuke Takeda, Akihiro Iida, Yu Otake, Sumihisa Orita, Seiji Ohtori

Disclosures

Human subjects: Informed consent for treatment and open access publication was obtained or waived by all participants in this study. Institutional Review Board of the School of Medicine, Chiba University issued approval 2805. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Balagué F, Mannion AF, Pellisé F, Cedraschi C: Non-specific low back pain. Lancet. 2012, 379:482-491. 10.1016/S0140-6736(11)60610-7
2. Prather H, van Dillen L: Links between the hip and the lumbar spine (hip-spine syndrome) as they relate to clinical decision making for patients with lumbopelvic pain. PM R. 2019, 11 Suppl 1:S64-72. 10.1002/pmrj.12187
3. Govil G, Tomar L, Dhawan P: Knee-spine syndrome: management dilemma when knee osteoarthritis coexists with spine degeneration. Cureus. 2022, 14:e24939. 10.7759/cureus.24939
4. Abbasi S, Mousavi SH, Khorramroo F: Association between lower limb alignment and low back pain: a systematic review with meta-analysis. PLoS One. 2024, 19:e0311480. 10.1371/journal.pone.0311480
5. Hafeez M, Tariq F, Sattar MI, Saeed A, Rahim S, Ghauri MW: Association of non-specific mechanical low back pain with limited ankle dorsiflexion among female students. J Riphah Coll Rehabil Sci. 2024, 12:183-6. 10.53389/JRCRS.2024120403
6. Almansoof HS, Nuhmani S, Muaidi Q: Role of ankle dorsiflexion in sports performance and injury risk: a narrative review. Electron J Gen Med. 2023, 20:521. 10.29333/ejgm/13412
7. Reis e Silva M, Lerebourg L: The effect of ankle dorsiflexion on sagittal posture and core muscle activation . Biomechanics. 2024, 4:812-26. 10.3390/biomechanics4040060
8. Bennell KL, Talbot RC, Wajswelner H, Techovanich W, Kelly DH, Hall AJ: Intra-rater and inter-rater reliability of a weight-bearing lunge measure of ankle dorsiflexion. Aust J Physiother. 1998, 44:175-80. 10.1016/S0004-9514(14)60377-9
9. Wiyanad A, Chokphukiao P, Suwannarat P, et al.: Is the occiput-wall distance valid and reliable to determine the presence of thoracic hyperkyphosis?. Musculoskelet Sci Pract. 2018, 38:63-8. 10.1016/j.msksp.2018.09.010
10. Faul F, Erdfelder E, Buchner A, Lang AG: Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009, 41:1149-60. 10.3758/BRM.41.4.1149
11. Park MW, Park SJ, Chung SG: Relationships between skeletal muscle mass, lumbar lordosis, and chronic low back pain in the elderly. Neurospine. 2023, 20:959-68. 10.14245/ns.2346494.247
12. Gerbershagen HJ, Rothaug J, Kalkman CJ, Meissner W: Determination of moderate-to-severe postoperative pain on the numeric rating scale: a cut-off point analysis applying four different methods. Br J Anaesth. 2011, 107:619-26. 10.1093/bja/aer195
13. Woo A, Lechner B, Fu T, et al.: Cut points for mild, moderate, and severe pain among cancer and non-cancer patients: a literature review. Ann Palliat Med. 2015, 4:176-83. 10.3978/j.issn.2224-5820.2015.09.04
14. Chu EC: Regional interdependence between senile kyphosis and cervical postural syndrome: a case report . AME Case Rep. 2025, 9:52. 10.21037/acr-24-41
15. Wertli MM, Eugster R, Held U, Steurer J, Kofmehl R, Weiser S: Catastrophizing-a prognostic factor for outcome in patients with low back pain: a systematic review. Spine J. 2014, 14:2639-57. 10.1016/j.spinee.2014.03.003
16. Pincus T, Burton AK, Vogel S, Field AP: A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine (Phila Pa 1976). 2002, 27:E109-20. 10.1097/00007632-200203010-00017