Developing a clinical prediction rule to identify patients with lumbar disc herniation who demonstrate short-term improvement with mechanical lumbar traction

Kazuya Hirayama, MHS, PT1,2, Eiki Tsushima, MSc, PhD, PT2, Hiroki Arihara, MHS, PT1 and Yoichi Omi, PhD, MD3

1) Department of Rehabilitation, Omi Orthopedic Clinic
2) Graduate School of Health Sciences, Hirosaki University
3) Department of Orthopedic Surgery, Omi Orthopedic Clinic

ABSTRACT. Objective: To develop a clinical prediction rule (CPR) that predicts treatment responses to mechanical lumbar traction (MLT) among patients with lumbar disc herniation (LDH). Method: This study was an uncontrolled prospective cohort study. The subjects included 103 patients diagnosed with LDH for which they underwent conservative therapy. The subjects received MLT for 2 weeks, and the application of any other medication was left at the discretion of the attending physician. The initial evaluation was performed prior to the initiation of treatment. The independent variables from the initial evaluation were imaging diagnosis, Oswestry Disability Index (ODI), Fear-Avoidance Beliefs Questionnaire score, visual analog scale, medical interview, physical examination. The patients whose ODI after 2 weeks of treatment improved by ≥50% of that at the initial evaluation were defined as responders. Results: Of the 103 subjects, 24 were responders, and the five predictors selected for the CPR were limited lumbar extension range of motion, low-level fear-avoidance beliefs regarding work, segmental hypomobility in the lumbar spine, short duration of symptoms, and sudden onset of symptoms. For the patients with at least three of the five predictors, the probability of their ODI greatly improving increased from 23.3% to 48.7% compared with the patients without these predictors (positive likelihood ratio, 3.13). Conclusion: Five factors were selected for the CPR to predict whether patients with LDH would demonstrate short-term improvement following conservative therapy with MLT.

Key words: lumbar disc herniation, mechanical lumbar traction, clinical prediction rule

Many attempts to identify effective interventions for patients with low back pain (LBP) have ended unsuccessfully1. Among the physical therapies for patients with LBP, lumbar traction therapy is one where its efficacy has been questioned. A systematic review2 of 32 randomized controlled trials revealed that traction only or a complex treatment including traction yielded only slight or no efficacy towards pain intensity, function, and ability to return to work. Even the Japanese clinical practice guideline of management of LBP3 state that traction has a low possibility of being effective for patients with LBP in general and there is much conflicting evidence for patients with sciatica, indicating that no fixed conclusion has been reached. Despite such, traction is still being practiced in many institutions; according to a survey4 on 121 medical institutions in Japan, the primary conservative therapy for patients with LBP is thermotherapy, followed by traction, which was performed in 24.5% of 4,598 patients with LBP.

Incidentally, the importance of the classification for performing better physical therapy management for patients with LBP is recently becoming recognized. The clinical prediction rule (CPR) that increases the precision of clinical
decision-making by combining clinical findings has been reported by several researchers as a method for performing a more effective classification. Some examples include lumbar manipulation CPR and lumbar stabilization exercise CPR. Notably, the CPR for predicting patients with LBP who demonstrate short-term improvement with mechanical lumbar traction (MLT) has been reported by Cai et al. According to their report, the percentage of short-term improvement due to MLT is said to increase from 19.4% to 69.2% in individuals who have the predictors of low-level fear-avoidance beliefs, no neurological deficit, age of ≥ 30 years, and non-manual worker compared with that in those who do not.

However, this CPR cannot be applied in the treatment provided in Japan. It is imperative that physical therapists in Japan, who lack a direct access system, perform treatment under the diagnosis and instructions of a physician. The reason for such is that the development and clinical application of the CPRs for patients with non-specific LBP, low back and leg pain in general are not as realistic in Japan as in foreign countries. Therefore, the application of the CPRs developed in countries other than Japan in the physical therapies performed in Japan remains controversial. If the development of a physical therapy-related CPR in the Japanese medical treatment system is envisaged, a CPR that is developed for a specific disease based on a physician’s diagnosis seems realistic.

There is also the opinion that the subgroup of patients with LBP who respond to traction consists of patients with radiculopathy in the acute phase and patients with neurological deficits or those who do not present the centralization phenomenon caused by lumbar movements. Traction may be potentially effective when radiculopathy or leg pain is also present rather than LBP only. As such, this suggests that there is a significance in identifying patients with lumbar disc herniation (LDH), which is a representative disease that expresses radiculopathy or leg pain, who respond to traction.

The objective of this study is to develop a CPR that predicts patients with LDH who would demonstrate short-term improvement with MLT in order to find out characteristics of patients with LDH have the effect of traction. In contrast to the CPRs related to physical therapies in foreign countries, the CPR in this study was developed considering not only the evaluation findings by physical therapists, but also the diagnosis of a physician and imaging findings.

Methods

1) Subjects

The subjects included 127 individuals diagnosed with LDH in an orthopedic surgery clinic for whom MLT was prescribed between November 10, 2014 and March 31, 2018. During the study period, 277 individuals were diagnosed with LDH. Of these, 46 who underwent surgical therapy, 85 who did not receive physical therapy, and 19 who met the exclusion criteria described below were excluded from the study; as such, 127 patients were finally included. The reason that 85 patients did not receive physical therapy was the physician did not give a prescription or was inconvenient for them, however the details were unknown. In addition, the subjects of this study overlapped with 66 those of our previous research. The study purpose and methods were adequately explained to the subjects according to the Helsinki Declaration principles; thereafter, the subjects provided their consent. This study was performed after obtaining an approval of the ethics committee of Hirosaki University School of Health Sciences (reference number 2014-019). The exclusion criteria were as follows: pregnancy, history of surgery in the lumbar spine, history of spinal fracture, and sacroiliac joint disorder. Sacroiliac joint disorder was judged when an individual showed positive findings for at least three of five tests (i.e., distraction test, compression test, thigh thrust test, sacral thrust test, and Gaenslen’s test). In fact, all 19 individuals who met the exclusion criteria had a history of surgery in the lumbar spine. Among the subjects, 24 for whom 2 weeks of follow-up from the initiation of the study could not be performed were excluded. Therefore, as shown in Table 1, 103 individuals were finally included as the subjects of the analysis (follow-up rate: 81.1%).

2) Measures

Prior to the initiation of treatment (initial evaluation), the imaging diagnosis by a physician and patient-reported outcome (PRO) were obtained, and a medical interview and examination were performed by a physical therapist. The list of the evaluation items is presented in Table 2.

For the imaging diagnosis, a spine specialist classified the level, direction (central/other), and type (non-contained/contained) of the hernia using magnetic resonance imaging (MRI). The transligamentous extrusion type and sequestrated type were categorized as the non-contained type, while the protrusion type and subligamentous extrusion type were classified as the contained type.

The PRO was assessed using a questionnaire on the visual analog scale (VAS) of pain, the Japanese version of the Oswestry Disability Index (ODI), and the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ). For the VAS, a range of 0-100 mm was set, with 0 mm representing no pain, and 100 mm representing maximum pain, and used to evaluate pain in the last few days of the study period. The Japanese version of the ODI demonstrated similar high reliability and validity as the original version of the ODI. The FABQ also demonstrates similar high reliability and validity as the original version. We did not use
Either one of two physical therapists performed the medical interview and physical examination. The details of the physical examination are presented below. The centralization and peripheralization phenomenon from the repeated movements test by McKenzie\(^{16}\) were evaluated. Active movements towards the flexion and extension directions were repeated 10 times in the standing position. “Centralize” was considered in individuals whose symptoms distribution converged towards the proximity of the lumbar area and “peripheralize” in those whose symptoms distribution scattered towards the periphery from the lumbar area. The inter-rater reliability (κ coefficient) of the centralization phenomenon was reported to be ≥0.7\(^{17}\). For lumbar range of motion, flexion, extension, and lateral bend were measured with the subjects in the standing position. Flexion/extension was measured in units of 0.1 cm using the modified Schöber test (MST)\(^{18}\). The examiner checked the lumbosacral junction while standing behind the subjects who were in the standing position. The examiner placed a mark at 10 cm above and another mark at 5 cm below the lumbosacral junction and measured the distance of the top and bottom marks at the time of maximum flexion and maximum extension using a tape measure. The reliability of the lumbar flexion/extension range of motion measured using the MST was reported to be high; even when the subjects were Japanese, the intra-rater intraclass correlation coefficient (ICC) was ≥0.8, and the inter-rater ICC was ≥0.7\(^{19}\). Lateral bend was measured in 1° units by the single inclinometer method\(^{20}\) using an inclinometer (Fabrication Enterprises, Bubble Inclinometer). For the straight leg raise (SLR) test, the hip was passively flexed until resistance was felt while the subjects were in the supine position with their knees kept extended. The SLR moveable range was measured in 1° units using the bubble inclinometer touching the tibial tuberosity. If sciatica occurred during the same procedure, a positive SLR test finding was judged. Hypomobility in any segment of the lumbar spine and pain during the tests were evaluated using the posterior-anterior mobility test (PAMT)\(^{21}\). For neurological deficit, superficial sensory testing, deep tendon reflex (patellar tendon and Achilles tendon) assessment, and manual muscle testing (anterior tibial muscle, extensor hallucis longus muscle, and peroneal muscle) were performed. When sensory disturbance, weakened tendon reflex, or decreased muscle strength was observed in the site corresponding to the hernia level, presence of a neurological deficit was judged.

### 3) Intervention

MLT was performed for 15 min using a lumbar traction device (Minato Medical Science, TC-30D) with the subjects in the semi-Fowler position and intermittent tractions applied (15 s of traction and 3 s of rest). The traction force was set at 30-40% of the subjects’ weight. The condition of the traction was determined in reference to the textbook. When the symptoms worsened during the procedure, the traction force was decreased. In addition, at the initiation of the procedure, the subjects were also instructed to maintain activities as much as possible and to move and keep a posture that does not place stress on the lumbar spine. The duration of the intervention was 2 weeks and was performed everyday provided that the subjects could tolerate it. The combined use of pain medication and spinal nerve root block (SRB) was left at the discretion of the physician.

The ODI was re-evaluated 2 weeks after the intervention was initiated, and the improvement rate was calculated. The formula for the improvement rate was as follows: \((\text{initial score-final score}) / \text{initial score} \times 100\). A subject with an improvement rate of ≥50% was considered as an MLT responder and that with a rate of <50% as an MLT non-responder. The validity of the threshold of ≥50% improvement in the ODI being used to define a successful outcome of LBP treatment was reported to be high\(^{22}\). For treatments administered from 2 weeks onwards, the treatment method was change as needed on the judgment of the physician.

### Table 1. Characteristics of subjects (N=103)

| Sex (N) | Men/Women | 60/43 |
|---------|-----------|------|
| Age (years) | 43.7±14.1 |
| Level of hernia (N) | L2/3 | 7 |
| | L3/4 | 6 |
| | L4/5 | 40 |
| | L5/S | 50 |
| Type of hernia (N) | Protrusion | 26 |
| | Subligamentous extrusion | 41 |
| | Transligamentous extrusion | 14 |
| | Sequestration | 22 |
| Direction of hernia (N) | Central/Other | 26/77 |

Mean ± Standard deviation (SD) unless otherwise indicated.
Table 2. Comparison of patients’ initial evaluation between the responder and non-responder groups

| Variables                        | Responder (N=24) | Non-responder (N=79) | P value | Test     |
|----------------------------------|------------------|----------------------|---------|----------|
| Sex (% women)                    | 45.8%            | 40.5%                | 0.643   | $\chi^2$ |
| Age (years)                      | 44.2±14.4        | 43.5±14.1            | 0.841   | t        |
| MRI findings                     |                  |                      |         |          |
| Level of hernia (%)              |                  |                      |         |          |
| L2/3                             | 12.5%            | 5.1%                 | 0.203   | Fisher   |
| L3/4                             | 8.3%             | 5.1%                 | 0.622   | Fisher   |
| L4/5                             | 29.2%            | 41.8%                | 0.267   | $\chi^2$ |
| L5/S                             | 45.8%            | 49.4%                | 0.762   | $\chi^2$ |
| Direction of hernia (%)          |                  |                      |         |          |
| Central type                     | 29.2%            | 24.1%                | 0.613   | $\chi^2$ |
| Type of hernia (%)               |                  |                      |         |          |
| Non-contained type               | 29.2%            | 36.7%                | 0.497   | $\chi^2$ |
| Treatment content                |                  |                      |         |          |
| Pain medication (%)              |                  |                      |         |          |
| Spinal nerve root block (%)      |                  |                      |         |          |
| Frequency of traction (times)    | 6.3±3.6          | 6.3±3.2              | 0.910   | U        |
| Patient reported outcome         |                  |                      |         |          |
| Pain ratings (VAS) (mm)          | 50.1±26.0        | 51.8±22.7            | 0.773   | U        |
| ODI score (%)                    | 34.5±21.6        | 27.3±12.3            | 0.331   | U        |
| FABQ physical activity subscale (points) | 12.5±5.4    | 14.2±4.8             | 0.203   | U        |
| FABQ work subscale (points)      | 13.3±9.4         | 17.0±9.8             | 0.108*  | U        |
| Medical interview                |                  |                      |         |          |
| Mode of onset (% sudden)         | 66.7%            | 46.8%                | 0.089*  | $\chi^2$ |
| Duration of symptoms (days)      | 34.0±79.9        | 92.6±249.4           | 0.055*  | U        |
| Manual worker (%)                | 50.0%            | 53.2%                | 0.786   | $\chi^2$ |
| Desk worker (%)                  | 25.0%            | 29.1%                | 0.695   | $\chi^2$ |
| Low back symptoms only (%)       | 37.5%            | 21.5%                | 0.115*  | $\chi^2$ |
| Buttock/thigh symptoms present (%) | 50.0%        | 68.4%                | 0.101*  | $\chi^2$ |
| Symptoms distal to knee (%)      | 45.8%            | 40.5%                | 0.643   | $\chi^2$ |
| Prior history of LBP (%)         | 87.5%            | 86.1%                | 1.000   | Fisher   |
| Walking ranked as worse position (%) | 25.0%  | 29.1%                | 0.695   | $\chi^2$ |
| Standing ranked as worse position (%) | 25.0%  | 34.2%                | 0.399   | $\chi^2$ |
| Sitting ranked as worse position (%) | 58.3%  | 62.0%                | 0.745   | $\chi^2$ |
| Walking ranked as better position (%) | 25.0%  | 31.6%                | 0.534   | $\chi^2$ |
| Standing ranked as better position (%) | 25.0%  | 31.6%                | 0.534   | $\chi^2$ |
| Sitting ranked as better position (%) | 12.5%  | 19.0%                | 0.555   | Fisher   |
| No position relieves pain (%)    | 58.3%            | 38.0%                | 0.077*  | $\chi^2$ |
| Physical examination             |                  |                      |         |          |
| Centralize with lumbar flexion (%) | 0%            | 1.3%                 | 1.000   | Fisher   |
| Centralize with lumbar extension (%) | 4.2%       | 13.9%                | 0.286   | Fisher   |
| Peripheralize with lumbar flexion (%) | 20.8%       | 30.4%                | 0.363   | $\chi^2$ |
| Peripheralize with lumbar extension (%) | 0%        | 8.9%                 | 0.196   | Fisher   |
| MST (range of lumbar flexion) (cm) | 17.9±1.7   | 18.5±1.5             | 0.285   | U        |
| MST (range of lumbar extension) (cm) | 13.5±0.6  | 13.1±0.7             | 0.041*  | U        |
| Left and right lateral bending discrepancy (°) | 3.1±2.7  | 3.8±3.1              | 0.284   | U        |
| Left and right SLR discrepancy (°) | 5.6±8.1    | 9.5±10.8             | 0.065*  | U        |
| Positive SLRT (%)                | 25.0%            | 26.6%                | 0.877   | $\chi^2$ |
| Hypomobility present with PAMT (%) | 20.8%    | 40.5%                | 0.079*  | $\chi^2$ |
| Pain with PAMT (%)               | 29.2%            | 29.1%                | 0.996   | $\chi^2$ |
| Neurological deficit involvement (%) | 12.5%  | 5.1%                 | 0.349   | Fisher   |

* p<0.15
Mean ± SD unless otherwise indicated.
$\chi^2$: chi-square test, t: two-sample t test, Fisher: Fisher’s exact test, U: Mann-Whitney’s U test
Figure 1. ROC curve of variables that were significantly different by bivariate analyses in continuous variables. AUC: area under the curve

Table 3. Predictors for the responder to MLT (multiple logistic regression analysis)

| Predictor                              | Odds ratio | 95% confidence interval | P value |
|----------------------------------------|------------|-------------------------|---------|
| MST extension ≥13.7 cm                 | 5.39       | 1.65-17.66              | 0.005   |
| FABQW ≤16                              | 4.74       | 1.40-16.07              | 0.012   |
| No hypomobility with PAMT              | 4.69       | 1.26-17.39              | 0.021   |
| Duration of symptoms ≤5 days           | 2.45       | 0.74-8.13               | 0.142   |
| Sudden onset                           | 2.34       | 0.73-7.53               | 0.155   |

Dependent variables: responder=1, non-responder=0

4) Data analysis

Bivariate analyses were performed for basic information, such as sex and age; MRI findings, such as level and type of the hernia; status of concomitant medication and SRB; frequency at which traction was performed; VAS and FABQ score; ODI score before the intervention; medical interview responses; and physical examination. The differences between the responder and non-responder were investigated. For continuous variables, the Shapiro-Wilk test was performed to confirm that a normal distribution was being followed. For p<0.05, the Mann-Whitney U-test was used, and for other p values, a two-sample t-test was used. For categorical variables, the χ² test or Fisher’s exact test was performed. Since these bivariate analyses were per-
formed to reduce the number of independent variables used in the subsequent multiple logistic regression analysis, the level of significance of these tests was set at \( p=0.15 \) as well as the previous CPR study\(^5\). Though there were methods of level of significance \( p=0.25, 0.20, 0.15 \) and 0.10 in previous studies\(^5-24\), \( p=0.15 \) was the smallest without variables being reduced too much. When a significant difference regarding the continuous variable was found, a receiver-operating characteristic (ROC) curve was drawn, and the sensitivity and specificity were calculated after obtaining the cut-off value. The score close to the left-most superior angle of the ROC curve was set as the cut-off value. For categorical variables, the sensitivity and specificity were calculated for each category. A stepwise method based on the Akaike’s Information Criterion was used in the multiple logistic regression analysis, and the level of significance was set at \( p=0.05 \). The multiple logistic regression analysis was carried out with responder/non-responder as dependent variable and significant variables by bivariate analyses as independent variables. Thereafter, the accuracy statistics commensurate to the number of positive items of the independent variables selected in the prediction model were calculated. For diagnostic accuracy, the sensitivity, specificity, and positive likelihood ratio (PLR) were used. The software used for all statistical analyses was R, 2.8.1, CRAN.

### Results

There were 24 (23.3%) responders and 79 (76.7%) non-responders. The bivariate analyses revealed that for nine variables, the \( p \) value was <0.15 (Table 2). The cut-off values for the FABQ work subscale, symptoms duration, lumbar extension range of motion, and difference in the right and left SLR angles, i.e., continuous variables, were 16, 5, 13.7, and 3 points (Fig. 1). These variables were also analyzed using a multiple logistic regression analysis; based on the results, the five items shown in Table 3 were selected (model \( \chi^2=23.37, p<0.001 \)). As the predictive model fitted the data (Hosmer-Lemeshow test, \( p=0.933 \)), the five selected items were set as the predictors of the CPR.

The correlation of the responders/non-responders with the number of predictors becoming positive is shown in Table 4. Five of 5 subjects with no variables present and 20 of 21 subjects with 1 of 5 variables present were in the non-responder group. Fourteen of 25 subjects with 3 of 5 variables present were in the responder. In addition, the accuracy statistics based on the number of positive items among the CPR predictors were calculated (Table 5). PLR was the highest in subjects with three or more variables present, furthermore it increased his or her probability of respond to MLT from 23.3% (24/103) to 48.7%.

### Discussion

MST is an evaluation of the range of motion of the overall lumbar spine, while PAMT is a segmental evaluation, as mentioned in method. MST extension was selected as one of the predictors of MLT responders and non-responders, and patients with limited lumbar extension range of motion were interpreted as ready responders. However, as the simultaneous presence of hypomobility in the PAMT also had an effect, these patients may also be considered to have no segmental hypomobility in the lumbar spine readily demonstrating improvement. Although these two predictors seem to be conflicting, patients with a narrow extension range of motion for the “overall” lumbar spine may be considered to respond readily to MLT and patients with “segmental” hypomobility to respond poorly to MLT. Several studies\(^25-26\) showed that traction could not affect on a specific segment but affect on multiple segments. Therefore Sugawara and Sakaguchi\(^27\) described that manual

---

**Table 4.** No. of subjects in the responder and non-responder groups at each level of CPR

| No. of predictor variables present | No. of subjects in responder group (N=24) | No. of subjects in non-responder group (N=79) |
|-----------------------------------|------------------------------------------|------------------------------------------|
| 0                                 | 0                                        | 5                                        |
| 1                                 | 1                                        | 20                                       |
| 2                                 | 4                                        | 34                                       |
| 3                                 | 14                                       | 11                                       |
| 4                                 | 5                                        | 8                                        |
| 5                                 | 0                                        | 1                                        |

**Table 5.** Accuracy statistics (with 95%CI) for each level of the prediction model

| No. of predictors present | Sensitivity | Specificity | PLR | Probability of responder to MLT |
|---------------------------|-------------|-------------|-----|---------------------------------|
| ≥1                        | 1.00 (1.00-1.00) | 0.06 (0.01-0.12) | 1.07 (1.01-1.13) | 24.5%   |
| ≥2                        | 0.96 (0.88-1.00) | 0.32 (0.21-0.42) | 1.40 (1.18-1.66) | 29.9%   |
| ≥3                        | 0.79 (0.63-0.95) | 0.75 (0.65-0.84) | 3.13 (2.05-4.78) | 48.7%   |
| ≥4                        | 0.21 (0.05-0.37) | 0.89 (0.82-0.96) | 1.83 (0.71-4.72) | 35.7%   |
| =5                        | 0.00 (0.00-0.00) | 0.99 (0.96-1.00) | 0.00 (0.00-0.00) | 0%      |
therapy was appropriate rather than traction when would like to mobilize a specific segment.

The patients with an FABQ work subscale score of ≤16 points were predicted to demonstrate improvement readily. This was a predictor backed even by previous studies. Fear-avoidance beliefs are a psychosocial factor of LBP. It is reported that the FABQ work subscale in particular is a prognostic factor of LBP and is one of the factors involved in enabling patients with sciatica to return to work. The cutoff value of FABQW in the acute and subacute LBP has reports of 20 in the factor predicting nonrecovery, and 14 in the CPR predicting days to recovery. The cutoff value of FABQW in this study was similar to the value to predict prognosis of the acute and subacute LBP.

For patients with a symptoms duration of ≤5 days, the fact that those with a sudden onset were predicted to demonstrate improvement readily shows that the effect of improvement in inflammation of the nerve root surroundings resulting from natural healing and medication, rather than the efficacy of MLT, is highly likely to be observed. MLT might be effective on acute pain with LDH, although the evidence to explain this mechanism are not obtained at present.

As shown in Table 5, the prediction accuracy was good when the standard was set for at least three predictors to be positive rather than for all five to be positive. The likelihood ratio is a standard value that demonstrates the usefulness of CPRs. Jaeschke et al. showed that a PLR of >2.0 indicates “small”; >5.0, “moderate”; and >10.0, “large” changes from pretest to posttest probability. When the number of positive factors was at least three, the PLR was 3.13, and the probability of 23.3% before the calculation changed to 48.7% after the calculation, which was a small change in probability. When the reference standard was set to ≥2 positive factors, the sensitivity was as high as 0.96. In other words, when the number of positive factors was below two, the percentage of patients falling under non-responders was high (25/26). This suggests that notable improvement cannot be expected with conservative therapy using MLT; this finding is then effective for excluding decision.

Herein, the level and type of the hernia, status of a neurological deficit, and changes of symptoms distribution with movements (centralization and peripheralization) were not extracted as significant predictors. The presence of a neurological deficit indicates that patients have a serious nerve root disorder, thus suggesting that the percentage of improvement by conservative therapy using MLT is likely low. However, in this study, the physician’s diagnosis was a prerequisite; therefore, there were only a few patients with serious nerve root disorder for which surgery was indicated. Fritz et al. stated that one of the characteristics of the subgroup of patients who benefit from traction is onset of the peripheralization phenomenon due to extension movements. However, among the subjects in this study, a few had an onset of centralization or peripheralization phenomenon, and a correlation could not be made.

This study has some limitations. First, it was a single-center survey study, and the characteristics of the subjects likely varied owing to the diagnosis of the physician, making the external validity unclear. Second, the sample size was small, according to the criteria (smaller number of outcome carrier or noncarriers > 10 × number of independent variables to use for a prediction model). Therefore, there are possibilities having low estimated precision about the variables that confidence interval of the odds ratio are large (duration of symptoms, sudden onset). Third, all the potential predictive variables have not necessarily been examined. Because we might have missed significant predictors other than we tested in this study, predictors should be validated and updated in the future. To solve these three limitations, the next steps that are required are to examine the external validity and to perform an impact study via an RCT. As fourth limitation, whether the traction conditions are appropriate is unclear. It has been reported that MLT where maximum traction force is used is not necessarily effective compared with sham traction. There is no consensus on the effective traction force as well as the traction and rest times.

Taking such limitations into consideration, immediate clinical application would be difficult. However, even among the same patients with LDH, there were those who demonstrated improvement following conservative therapy using MLT and those who did not. It became evident that the percentage of patients demonstrating a notable improvement (ODI of ≥50%) was low, not exceeding 23.3%. This shows that the efficacy of lumbar traction is limited and that establishing such indications is important.

**Conclusion**

Five factors were selected for the CPR for predicting short-term improvement after conservative therapy using MLT among the patients with LDH. The probability of improvement increased when a patient had at least three of the five predictors; the probability of no improvement was high when a patient had less than two predictors. We hope this CPR becomes the useful tool for clinical decision making of providing appropriate treatment to LDH patients by examining external validity in the future.

**Conflict of Interest:** The authors have no financial relationships to disclose.

**References**

1) van Tulder MW, Koes BW, et al.: Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common inter-
ventions. Spine. 1997; 22: 2128-2156.

2) Wegner I, Widyahening IS, et al.: Traction for low-back pain with or without sciatica. Cochrane Database Syst Rev. 2013 Aug 19; 8: CD003010. doi: doi: 10.1002/14651858.CD003010. pub 5.

3) The Japanese Orthopaedic Association: The Japanese Orthopaedic Association clinical practice guideline for management of low back pain. Nankodo, Tokyo, 2012, pp. 46-47 [In Japanese].

4) Ebata S, Oda H, et al.: Factual survey of patients at conservative treatment outpatient clinics; evaluation of 4,598 cases at 121 facilities in Saitama prefecture. Orthopedic Surgery. 2011; 62: 101-107 [In Japanese].

5) Flynn T, Fritz J, et al.: Clinical prediction rule for classifying patients with low back pain likely to respond to a manipulation technique. Spine. 2002; 27: 2835-2843.

6) Hicks GE, Fritz JM, et al.: Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. Arch Phys Med Rehabil. 2005; 86: 1753-1762.

7) Cai C, Pua YH, et al.: A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with mechanical lumbar traction. Eur Spine J. 2009; 18: 554-561.

8) Krause M, Refshauge KM, et al.: Lumbar spine traction: evaluation of effects and recommended application for treatment. Man Ther. 2000; 5: 72-81.

9) Fritz JM, Cleland JA, et al.: Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. J Orthop Sports Phys Ther. 2007; 37: 290-302.

10) Hirayama K, Tsushima E, et al.: External validation of the clinical prediction rule for lumbar traction. Rigakuryohokenkyu. 2017; 34: 20-24 [In Japanese].

11) Laslett M, Young SB, et al.: Diagnosing painful sacroiliac joints: A validity study of a McKenzie evaluation and sacroiliac provocation tests. Aust J Physiother. 2003; 49: 89-97.

12) Fujiwara A, Kobayashi N, et al.: Association of the Japanese Orthopaedic Association score with the Oswestry Disability Index, Roland-Morris Disability Questionnaire, and short-form 36. Spine. 2003; 28: 1601-1607.

13) Matsuura K, Kikuchi N, et al.: Psychometric properties of the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ). J Orthop Sci. 2014; 19: 26-32.

14) Fairbank JC, Couper J, et al.: The Oswestry low back pain disability questionnaire. Physiotherapy. 1980; 66: 271-273.

15) Waddell G, Newton M, et al.: A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain. 1993; 52: 157-168.

16) McKenzie RA: The Lumbar Spine: Mechanical Diagnosis and Therapy, Spinal Publications New Zealand, Waikanae, 1989.

17) Kilpikoski S, Airaksinen O, et al.: Interexaminer reliability of low back pain assessment using the McKenzie method. Spine. 2002; 27: 207-214.

18) Moll JM and Wright V: Normal range of spinal mobility. An objective clinical study. Ann Rheum Dis. 1971; 30: 381-386.

19) Umeno Y, Ishida K, et al.: Reliability and validity of Modified Schöber test for measuring lumbar spine motion. Journal of Hokkaido Physical Therapy. 2012; 29: 24-28 [In Japanese].

20) Waddell G, Somerville D, et al.: Objective clinical evaluation of physical impairment in chronic low back pain. Spine. 1992; 17: 617-628.

21) Maher CG, Latimer J, et al.: An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. Phys Ther. 1998; 78: 829-837.

22) Yano Y: Traction therapy. Chap 2. In: Amimoto K (ed): Hyourikaku, 2nd ed, Igakusyoin, Tokyo, 2004, pp. 26-28 [In Japanese].

23) Fritz JM, Hebert J, et al.: Beyond minimally important change: defining a successful outcome of physical therapy for patients with low back pain. Spine. 2009; 34: 2803-2809.

24) Katz MH: Multivariable analysis: A practical guide for clinicians, Cambridge University Press, Cambridge, 2006.

25) Reilly JP, Gersten JW, et al.: Effect of pelvic-femoral position on vertebral separation produced by lumbar traction. Phys Ther. 1979; 59: 282-286.

26) Chow DHK, Yuen EMK, et al.: Mechanical effects of traction on lumbar intervertebral discs: A magnetic resonance imaging study. Musculoskelet Sci Pract. 2017; 29: 78-83.

27) Sugawara H and Sakaguchi M: Effects and problem of traction therapy. The Journal of Japanese Physical Therapy Association. 2005; 32: 253-257 [In Japanese].

28) Werti MM, Rasmussen-Barr E, et al.: The role of fear avoidance beliefs as a prognostic factor for outcome in patients with nonspecific low back pain: a systematic review. Spine J. 2014; 14: 816-836.

29) Gurol L, Haugen AJ, et al.: Prognostic factors for return to work in patients with sciatica. Spine J. 2013; 13: 1849-1857.

30) Jaeschke R, Guyatt G, et al.: Users’ guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA. 1994; 271: 703-707.

31) Fritz JM, Lindsay W, et al.: Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. Spine. 2007; 32: 793-800.

32) Peduzzi P, Concato J, et al.: A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996; 49: 1373-1379.

33) McGinn TG, Guyatt GH, et al.: Users’ guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. JAMA. 2000; 284: 79-84.

34) Beurskens AJ, de Vet HC, et al.: Efficacy of traction for non-specific low back pain: a randomised clinical trial. Lancet. 1995; 346: 1596-1600.