Identifying Acute Lumbar Spondylolysis in Young Athletes with Low Back Pain: Retrospective Classification and Regression Tree Analysis

Masashi Aoyagi, MPhty,[a,b] Kei Naito, RPT,[a] Yuichi Sato, RPT, [a] Atsushi Kobayashi, MD, [a] Masaaki Sakamoto, PhD, [b] Steve Tumilty, PhD [c]

From the [a]Forest Orthopaedic Sports Clinic, Japan, [b]Graduate School of Health Sciences, Gunma University, Japan, and [c]School of Physiotherapy, University of Otago, New Zealand.

Acknowledgement date: July 7,2020. First revision date: October 6,2020. Acceptance date: November 12, 2020.

Study Design. Case control study

Objective. To establish an algorithm to distinguish acute lumbar spondylolysis (LS) from non-specific low back pain (NSLBP) among patients in junior high school by classification and regression tree (CART) analysis.

Background. Rapid identification of acute LS is important because delayed diagnosis may result in pseudarthrosis in the pars interarticularis. To diagnose acute LS, magnetic resonance imaging (MRI) or computed tomography is necessary. However, not all adolescent patients with low back pain (LBP) can access these technologies. Therefore, a clinical algorithm that can detect acute LS is needed.

Methods. The medical records of 223 junior high school-aged patients with diagnosed acute NSLBP or LS verified by MRI were reviewed. A total of 200 patients were examined for establishing the algorithm and 23 were employed for testing the performance of the algorithm. CART analysis was applied to establish the algorithm using the following data; age, gender, school grades, days after symptom onset, past history of LBP, days of past LBP, height, passive straight leg raising test results, hours per week spent in sports activities, existence of spina bifida, lumbar lordosis angle, and lumbosacral joint angle. Sensitivity and specificity of the algorithm and the area under the ROC curve were calculated to assess algorithm performance.

Results. The algorithm revealed that gender, days after symptom onset, days of past LBP, hours per week spent in sports activities, and existence of spina bifida were key predictors for identifying acute LS versus NSLBP. Algorithm sensitivity was 0.64, specificity was 0.92, and the area under the ROC curve was 0.79.

Conclusion. The algorithm can be used in clinical practice to distinguish acute LS from NSLBP in junior high school athletes, although referral to MRI may be necessary for definitive diagnosis considering the algorithm's sensitivity.

Keywords. Low Back Pain, Athletic Injuries, Clinical Decision-Making, Adolescent, Youth Sports, Clinical Prediction Rules, Algorithm, Stress Fracture, Diagnosis, Machine Learning

Level of Evidence. 4

Lumbar spondylolysis (LS) is a bony defect in the pars interarticularis of the lumbar vertebra that frequently seen in active adolescents with the incidence of 20% to 63% in some sports.[1,2] LS requires conservative or surgical treatment, resulting in extensive time away from sports activities. In conservative treatment, early diagnosis is important because acute LS can be united by an appropriate rehabilitation program and braces, while delayed diagnosis may result in chronic non-union of the pars fracture.[1]

Several risk factors for developing LS include innate factors such as the presence of spina bifida occulta,[3] increased lumbar lordosis,[4,5] and increased anterior pelvic tilt.[4,6] Men are more likely to suffer from LS than women.[3] Research has suggested that tightness of hamstrings and hip flexors could be related to LS development because this can increase stress in the lumbar spine by preventing movement of the hip joint and pelvis.[1,6,7] In terms of physique, Yanagisawa et al[8] found that patients with LS were taller than average for any particular age group; however, Wren et al[5] found no significant differences in height and weight between adolescents with and without LS. Furthermore, considering LS is a stress fracture, the amount of sports participation should be taken into consideration for LS development.

Differential diagnosis of acute LS from non-specific LBP (NSLBP) is challenging. Although there are several clinical tests for detecting this injury, these tests and test batteries cannot identify acute LS from other causes of LBP, raising the question of their diagnostic value.[9-11] Thus, imaging techniques are recommended. Plain radiography is not sensitive enough to detect acute LS.[9] Therefore, magnetic resonance imaging (MRI) and computed tomography (CT) are regarded as the gold standard for definitive diagnosis.[6] However, not all patients can access these advanced imaging techniques because of prohibitively expensive equipment and medical costs. To date, there is little research aimed at detecting acute LS with the combination of clinical findings. Thus, developing an algorithm to identify this injury that can be used in a clinical setting is important as it could aid diagnosis and referral for appropriate imaging investigation.

As the incidence of acute LS is complex, a classification and regression tree (CART) analysis was employed to identify patients with this injury. CART analysis is a nonparametric and multivariate statistical method to develop screening and diagnostic algorithms to identify risk factors for specific conditions.[12-14] It produces a tree diagram by binary division of the population with cutoff points and is easy to understand and use in clinical practice. The purpose of this study was to establish an algorithm using CART analysis to distinguish acute LS from NSLBP in young athletes. For the purpose of this project, acute LS was defined as causing pain leading to the initial diagnosis of LS.

MATERIALS AND METHODS

Subjects

This study retrospectively examined the medical and physical therapy records of patients who complained of LBP at Forest Orthopaedic & Sports Clinic from June 1, 2015, to August 31, 2019. Records included interview sheets and X-ray and MRI results. Inclusion criteria were junior high school students aged between 12 and 15,

participating in sports activities, diagnosed with acute NSLBP or LS by MRI, and received physical therapy treatment. As it was difficult to identify when stress fracture occurred, diagnosis of acute LS was confirmed by a high-intensity signal at the pars interarticularis in short T1 inversion recovery regardless of duration of the pain. NSLBP was diagnosed when no remarkable changes were found in the X-ray and MRI. Exclusion criteria were diagnosed with any other LBP-related injuries or diseases such as progressive LS, lumbar disc herniation, or lumbar disc disease and recurrence of acute LS (if patients presented to the clinic for treatment for acute LS more than once, only data from the first occurrence was included). Progressive LS was diagnosed when MRI did not reveal high-intensity signals and X-ray showed the existence of LS. Although sample size calculation was not conducted as CART analysis did not investigate differences between groups, about 200 participants were expected to obtain reliable results.[13,14]

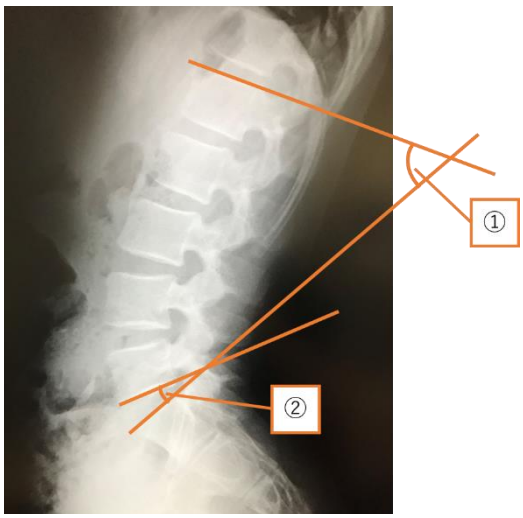

Figure 1. Lumbar lordosis angle and lumbosacral joint angle.[16]
① Lumbar lordosis angle: between upper endplate of the first lumbar vertebra and upper endplate of the sacrum. ② Lumbosacral joint angle: between lower endplate of the fifth lumbar vertebra and upper endplate of the sacrum.

Procedures and data collection

Data from the initial evaluation were retrospectively collected. These were; 1) personal information including age, gender, and grades in school, 2) injury information including days after symptom onset, past history of LBP, and days of past LBP, 3) physical characteristics including height and passive straight leg raising test results,[15] 4) sports activity information including type of sports activity and average hours per week spent on sports activities, and 5) finding on X-ray photography including existence of spina bifida, lumbar lordosis angle, and lumbosacral joint angle (Figure 1).[16] As most of acute LS occurred by repetitive microtrauma,[6] duration of the injury was not clearly identified. Therefore, we defined symptom onset as when chief complaint of LBP was noticed or when LBP was exacerbated resulting in seeking medical support. Once LBP had settled and the patient participated in sports activity, it was regarded as past LBP. Existence of spina bifida was measured by standing anteroposterior lumbar spine radiographs, while lumbar lordosis angle and lumbosacral joint angle were measured by lateral standing lumbar spine radiographs by ViewSend RAD (ViewSend ICT Co., Ltd., Japan). Subjective assessment was conducted by medical consultation and original interview sheet. The study protocol was approved by the Gunma University Ethics Committee (HS2019-180). Instead of obtaining informed consent, the opt-out method was employed as this study was retrospective and without any interventions.

Statistical analysis

All variables were selected based on the consideration of risk factors for LS discussed above

and could be gathered retrospectively.[1,3-8] Stochastic regression imputation was employed for missing data. CART analysis was used to establish an algorithm to distinguish acute LS from NSLBP. CART analysis can use both qualitative and quantitative data and does not require variable selection before building the algorithm as it selects the most important variables.[12] One potential disadvantage of the analysis is overfitting to the population used to create the model, which can prevent generalization.[17] As overfitting can occur when the number of nodes increases, we regulated the number of nodes by pruning with a complexity parameter. Sensitivity, specificity, accuracy, and the area under the receiver-operating characteristic (ROC) curve were calculated to measure algorithm performance. The performance was judged according to previous criteria (0.7 to 0.8: acceptable, 0.8 to 0.9: excellent, ≧ 0.9: outstanding).[18] For these statistical procedures, the full dataset was randomly divided into a trial dataset for building the algorithm (90% of the population) and a validation dataset for testing the performance of the algorithm (10% of the population). Prevalence ratios were calculated in each node to measure the strength of individual predictors.[14] As a sensitivity analysis, complete-case analysis was conducted with the full dataset. Descriptive statistics were used to describe the characteristics of the trial dataset. All analyses were conducted with R version 3.6.0. (see document, Supplemental Digital Content 1, which demonstrates packages and function used in R).

RESULTS

Overall, 223 patients were included. Of these, 124 patients were diagnosed with acute LS and 99 with NSLBP. There were no refusals in 6 months of opt-out. Of 223 participants, 200 were randomly selected as the trial dataset to establish the CART algorithm using the function of "sample" in R (see document, Supplemental Digital Content 1, which demonstrates this function in R), and 23 participants were used as the validation dataset. Table 1 shows the descriptive statistics of the trial dataset. Table 2 presents sports activities in which participants engaged. There were more male patients with acute LS than females (male to female ratio, 6.5:1), whereas the number of male and female patients with NSLBP was almost equal. Type of sports activity was similar in both groups. Many participants with acute LS played baseball, soccer, or basketball, while those with NSLBP mostly played soccer, basketball, or volleyball.

To regulate overfitting, nodes were pruned with a complexity parameter of 0.018 for reducing possible cross-validated error (Figure 2). The algorithm revealed that gender, days after symptom onset, days of past LBP, hours per week spent in sports activities, and existence of spina bifida were predictors for classifying acute NSLBP and LS. The sensitivity, specificity, and accuracy of the algorithm were 0.64, 0.92, and 0.78, respectively. The area under the ROC curve was 0.79, suggesting that the performance of the algorithm was acceptable.[18] Table 3 shows the prevalence ratios of each terminal node of the algorithm. The associations among predictors of node 10, 11, and 15 were statistically significant with the presence of acute NSLBP or LS.

Sensitivity analysis revealed that the first, second, and third predictors were the same for the sensitivity analysis and the main analysis, although cutoff points were slightly different.

Table1. Demographic characteristics for the trial data

Variable	Range	Trial Data (n=200)	Lumbar Spondylolysis (n=112)	Non-Specific Low Back Pain (n=88)
Gender		Male: 144 Female: 56	Male: 97 Female: 15	Male: 47 Female: 41
Grades		Grade7: 55 Grade8: 103 Grade9: 42	Grade7: 26 Grade8: 62 Grade9: 24	Grade7: 29 Grade8: 41 Grade9: 18
Age [year]	12 - 15	13.5 ± 0.8	13.6 ± 0.8	13.5 ± 0.7
Height [cm]	140 - 180	161.9 ± 7.7	163.4 ± 7.3	159.9 ± 7.7
Days after symptom onset [day]	0 - 180	20.2 ± 28.2	15.4 ± 18.7	26.3 ± 35.9
Past history of low back pain		67 (34%)	32 (29%)	35 (40%)
Days of past low back pain [month]	0 - 48	2.9 ± 6.8	2.2 ± 5.9	3.7 ± 7.7
Passive SLR test R [°]	40 - 100	62.8 ± 10.7	61.5 ± 10.3	64.4 ± 10.9
Passive SLR test L [°]	30 - 100	62.7 ± 10.7	61.7 ± 10.5	64.0 ± 10.8
Hours per week spent in sports activities [hours/week]	6 - 41	16.4 ± 6.3	16.8 ± 6.7	15.8 ± 5.7
Spina bifida		31 (16%)	12 (11%)	19 (22%)
Lumbar lordosis angle [°]	23.2 - 80.0	50.5 ± 9.1	51.6 ± 9.4	49.0 ± 8.6
Lumbosacral joint angle [°]	0.6 - 26.8	12.7 ± 5.6	12.2 ± 5.6	13.2 ± 5.6

Abbreviation: SLR; straight leg raising

Predictors after the fourth were different between the two analyses (Figure 3).

DISCUSSION

Identifying acute LS from NSLBP without MRI or CT is challenging because both show similar symptoms and signs in clinical examination. Using CART analysis, this study developed an algorithm based on various potential risk factors for LS to distinguish acute LS from NSLBP. Interactions between gender, days after symptom onset, days of past LBP, hours per week spent in sports activities, and the existence of spina bifida could identify acute LS in junior high school patients. Furthermore, based on the prevalence ratios of each terminal node, three factors were selected as strong predictors. Node 11 revealed that male patients who visited a medical facility within 32 days of symptom onset, had LBP within 10 months prior to onset or no past history of LBP, and participated in sports activity more than 17 hours per week were most likely to suffer from acute LS. Node 15 showed that in the above patients engaged in sports activity less than 17 hours per week, if spina bifida did not exist, they were likely to suffer from acute LS. In contrast, node 10 showed that female patients who played sports between 14 and 24 hours per week were most likely to have NSLBP.

Table2. Sports activities in which participants engaged

Sports activities	Lumbar Spondylolysis (n=112)	Non-Specific Low Back Pain (n=88)
Baseball	25	6
Soccer	24	18
Basketball	13	15
Volleyball	5	11
Tennis	5	8
T&F Sprint	6	3
Middle distance	0	2
Long distance	4	1
Jump	3	1
Others	4	3
Rugby	4	2
Softball	3	2
Gymnastics	2	4
Swimming	0	4
Other sports	14	8

Abbreviation: T&F; track and field

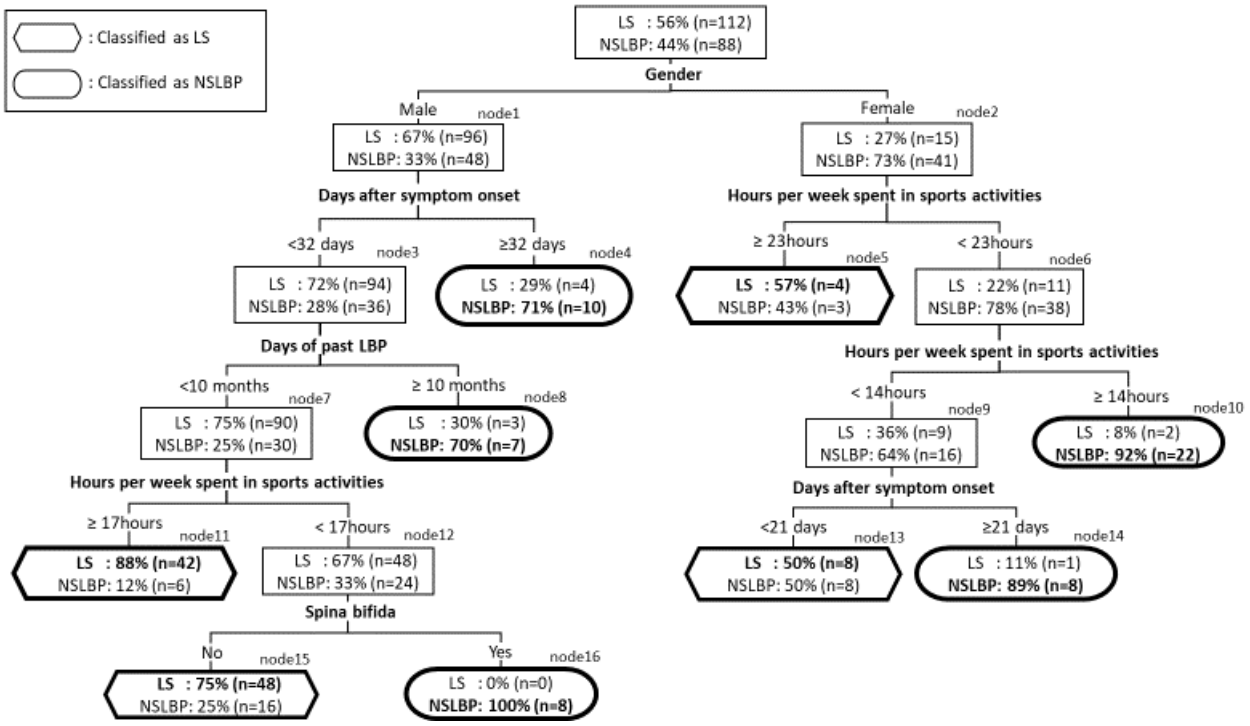

Figure 2. CART algorithm for identifying acute LS from NSLBP.
Terminal nodes are shown with bold frames. Hexagonal frames represent patients classified with LS while ellipse frames represent those classified with NSLBP. Incident of LS or NSLBP was shown in percentage term in each node. The patients in node 5, 11, 13, and 15 were classified into LS, while those in node 4, 8, 10, 14, and16 were classified into NSLBP.
Abbreviations: N; Number of patients, LBP; Low back pain, LS; Lumbar spondylolysis, NSLBP; Non-specific low back pain.

Table 3. Prevalence ratios in each terminal node

Terminal node	Prevalence ratio (95%CI)
4	0.49 (0.21 - 1.14)
5	1.02 (0.53 - 1.96)
8	0.52 (0.20 - 1.36)
10	0.13 (0.04 - 0.51)*
11	1.90 (1.55 - 2.33)*
13	0.88 (0.53 - 1.47)
14	0.19 (0.03 - 1.22)
15	1.59 (1.27 - 2.00)*
16**	-

*Statistically significant. ** In terminal node 16, prevalence ratio was not able to calculate as there was no LS patients.

Of all interactions in the algorithm, these three interactions could be particularly useful in clinical practice to differentiate acute LS from NSLBP.

Although the performance of this algorithm in identifying acute LS was acceptable based on the area under the ROC curve,[18] employing it in clinical setting comes with caveats. The specificity of the algorithm was excellent (0.92), but sensitivity was 0.64, suggesting that the model could produce false negatives. Therefore, referral to MRI might be necessary, especially for patients not classified as having acute LS. In contrast, for patients classified with acute LS, additional examination might be unnecessary due to the high specificity of the algorithm. Further study exploring how to improve the sensitivity is necessary before advocating for its use in clinical practice.

The first predictor was gender, suggesting that male adolescent patients were more likely to suffer from acute LS than female adolescents. A similar result was reported in previous research, although the proportion of males in this study was much higher (male to female ratio of 6.5:1 in this study compared with 2:1[3]). In male patients, days after symptom onset was related to the presence of acute

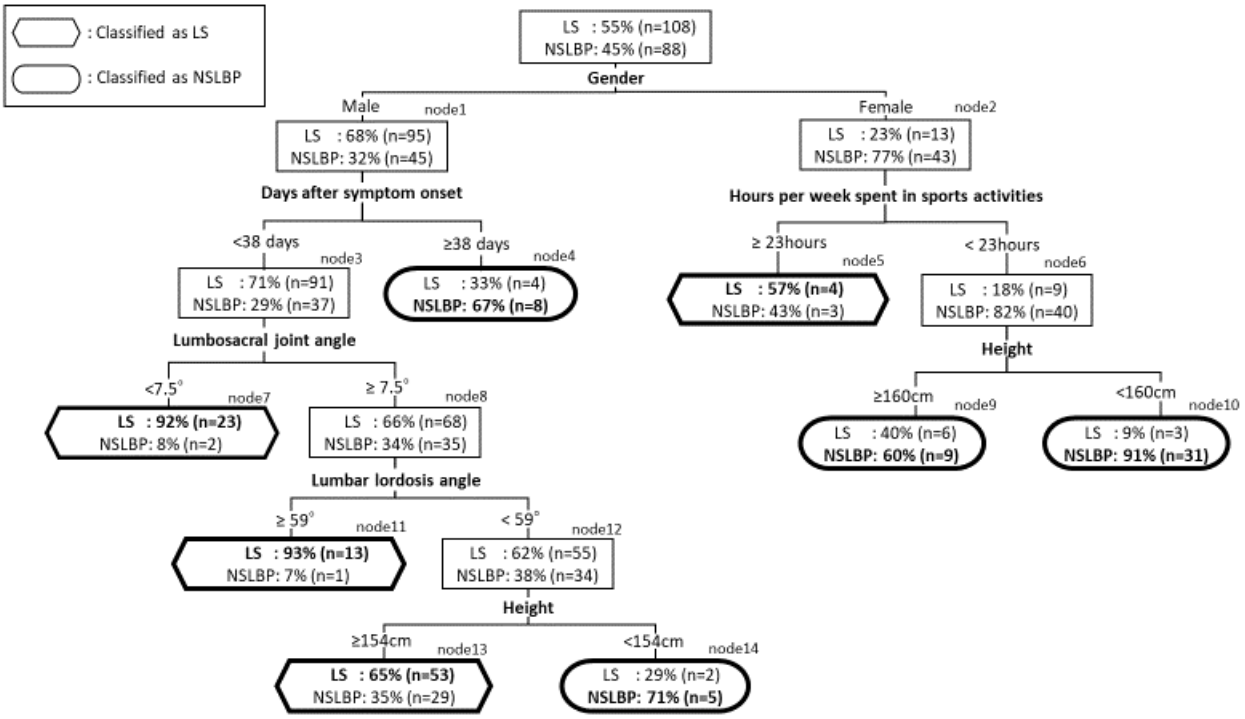

Figure 3. Sensitivity analysis (complete-case analysis with the total dataset)
Terminal nodes are shown with bold frames. Hexagonal frames represent patients classified with LS, while elliptical frames represent those classified with NSLBP. Incident of LS or NSLBP was shown in percentage term in each node. The patients in node 5, 7 11, and 13 were classified into LS, while those in node 4, 9, 10, and 14 were classified into NSLBP.
Abbreviations: N; Number of patients, LS; Lumbar spondylolysis, NSLBP; Non-specific low back pain.

LS. Those who sought medical attention within 32 days after onset were more likely to have acute LS than those presenting later. Thus, patients with this injury may feel more intense pain and/or dysfunction, resulting in them seeking medical treatment sooner. Alternatively, it can be postulated that those seeking help later were diagnosed with NSLBP because inflammation at the pars interarticularis is likely to have settled after 32 days. This is possible if patients ceased sports activities. However, in the author's experience, this is unrealistic because patients tend to visit medical facilities when they feel pain or their pain worsens and not after the pain decreases. Although it is also possible for acute LS to advance to the progressive phase after 32 days, this was not frequent in our sample group as progressive LS was detected by oblique radiographs of the lumbar spine.

In male patients who visited medical facilities within 32 days after onset, days of past LBP was the third predictor. Those who experienced LBP <10 months from symptom onset or had not experienced LBP tended to be diagnosed with acute LS compared with those who experienced LBP ≥10 months before onset. The explanation for this might be that significant pain or short-term recurrence could drive patients to visit medical facilities sooner. Additionally, it may indicate that because stress on the pars interarticularis accumulates before it resolves, acute LS was likely to occur in adolescent patients who experienced repeated LBP over a period of 10 months. Longer periods of LBP recurrence may be a characteristic of NSLBP in the current population.

Hours of sports participation per week was a predictor for both male and female patients. In males, ≥17 hours of sports activity per week was significantly associated with the presence of acute LS. In females, between 14 and 23 hours of sports participation was related to NSLBP, while ≥23 hours of participation could relate to an increase in acute LS occurrence. These associations support the idea that one of the causes of acute LS in young athletes is unaccustomed overload and that the amount of time spent participating in sports is an important factor to identify this injury.

Interestingly, in male patients, existence of spina bifida was inversely related to the presence of acute LS. In contrast, a previous study reported a positive relationship between spina bifida occulta and LS occurrence.[3] Although ethnicity in both studies was similar, the previous study examined patients aged 20–92 years,[3] whereas the current study focused on patients aged 12–15 years. Further research is necessary to confirm this difference as the possibility of overfitting to the population in our CART analysis cannot be disregarded.

The literature indicates that excessive lumbar lordosis, increasing anterior pelvic tilt, and hamstring tightness are thought to be risk factors for developing LS.[6] However, the current results demonstrated that physical factors including physique, alignment of the lumbar spine, and muscle flexibility were not strong indicators for distinguishing acute LS from NSLBP. As most of the predictors in the algorithm can be gathered in a subjective examination, medical consultation becomes much more important than physical examination in diagnosing acute LS among junior athletes with LBP. Furthermore, although many researchers have pointed out that sports involving repeated extension and rotation movement tends to cause LS,[6,19] the current results did not support this as many patients with NSLBP participated in such sports activities, including soccer, basketball, and volleyball.

There are some limitations in this study. First, due to the nature of CART analysis, the results might overfit to the trial population, which could prevent the algorithm from being generalized, even though the nodes were pruned. Second, sensitivity analysis showed that internal validity of the algorithm was questionable. The difference between the sensitivity and main analyses might be a result of the imputation methodology used. Multiple imputation is recommended because it can provide unbiased and valid estimates,[20] but this study employed a single imputation, which might lack precision, due to constraints on the project. Third, there might be selection bias as not all adolescent patients with LBP who presented to the clinic underwent MRI, despite this being recommended to the majority unless their symptoms were minor. However, considering that the difference in the number of patients with acute LS versus NSLBP was not large, the authors believe that such bias was minor. Based on these limitations, the current result should be interpreted with caution when used in clinical practice.

CONCLUSION

The CART analysis was used to establish an algorithm to identify acute LS from NSLBP in junior athletes without MRI or CT. According to the algorithm, gender, days after symptom onset, days of past LBP, hours per week spent in sports

activities, and the existence of spina bifida should be considered when acute LS is suspected. Clinicians may apply the algorithm to screen acute LS in junior high school patients with LBP and refer for appropriate imaging investigation for definitive diagnosis. However, as it could generate false negatives, it should be used with caution. For full clinical implications, the algorithm should be validated in different populations and refined with other possible identifiable factors of acute LS to improve discriminatory capacity in future study.

Key Points

- Early diagnosis of acute lumbar spondylolysis without MRI or CT is challenging although it is critical for bone union in adolescents.
- A clinical algorithm composed of gender, days after symptom onset, days of past LBP, hours per week spent in sports activities, and existence of spina bifida could distinguish acute LS from NSLBP.
- Although the performance of the algorithm was acceptable for clinical use (the area under the ROC curve was 0.79), referral to MRI may be needed to decrease the possibility of missing acute LS.

References

1) Leone A, Cianfoni A, Cerase A, et al. Lumbar spondylolysis: a review. *Skeletal Radiol*. 2011; 40: 683-700.
2) Stanitski CL. Spondylolysis and spondylolisthesis in athletes. *Oper Tech Sports Med*. 2006; 14: 141-146.
3) Sakai T, Sairyo K, Takao S, et al. Incidence of lumbar spondylolysis in the general population in Japan based on multi-detector computed tomography scans from two thousand subjects. *Spine (Phila Pa 1976)*. 2009; 34: 2346-2350.
4) Yin J, Peng BG, Li YC, et al. Differences of sagittal lumbosacral parameters between patients with lumbar spondylolysis and normal adults. *Chin Med J*. 2016; 129: 1166-1170.
5) Wren TAL, Ponrartana S, Aggabao PC, et al. Increased lumbar lordosis and smaller vertebral cross-sectional area are associated with spondylolysis. *Spine (Phila Pa 1976)*. 2018; 43: 833-838.
6) Lawrence KJ, Elser T, Stromberg R. Lumbar spondylolysis in the adolescent athlete. *Phys Ther Sport*. 2016; 20: 56-60.
7) Thein-Nissenbaum J, Boissonnault WG. Differential diagnosis of spondylolysis in a patient with chronic low back pain. *J Orthop Sports Phys Ther*. 2005; 35: 319-326.
8) Yanagisawa R, Tsukagoshi Y, Nakashima K, et al. Analysis of the physique of adolescents with lumbar spondylolysis. *The journal of clinical of sports medicine*. 2018: 26: 242-246. (Japanese)
9) Kobayashi A, Kobayashi T, Kato K, et al. Diagnosis of radiographically occult lumbar spondylolysis in young athletes by magnetic resonance imaging. *Am J Sports Med*. 2013; 41: 169-176.
10) Alqarni AM, Schneiders AG, Cook CE, et al. Clinical tests to diagnose lumbar spondylolysis and spondylolisthesis: a systematic review. *Phys Ther Sport*. 2015; 16: 268-275.
11) Sundell CG, Jonsson H, Adin L, et al. Clinical examination, spondylolysis and adolescent athletes. *Int J Sports Med*. 2013; 34: 263-267.
12) The SK, Zheng W, Ho KY, et al: Diagnosis of gastric cancer using near-infrared raman spectroscopy and classification and regression tree techniques. *J Biomed Opt*. 2008; 13: 1-8.
13) Bittencourt NF, Ocarino JM, Mendonça LD, et al. Foot and hip contributions to high frontal plane knee projection angle in athletes: a classification and regression tree approach. *J Orthop Sports Phys Ther*. 2012; 42: 996-1004.
14) Mendonça LD, Ocarino JM, Bittencourt NFN, et al. Association of hip and foot factors with patellar tendinopathy (jumper's knee) in athletes. *J Orthop Sports Phys Ther*. 2018; 48: 676-684.
15) Majlesi J, Togay H, Unalan H, et al. The sensitivity and specificity of the slump and the straight leg raising tests in patients with lumbar disc herniation. *J Clin Rheumatol*. 2008; 14: 87-91.
16) Caglayan M, Tacar O, Demirant A, et al. Effects of lumbosacral angles on development of low back pain. *J Musculoskelet Pain*. 2014; 22: 251-255.
17) Schaffer C. Overfitting avoidance as bias. *Machine Learning*. 1993; 10: 153-178.
18) Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010; 5: 1315-1316.
19) Sairyo K, Katoh S, Komatsubara S, et al. Spondylolysis fracture angle in children and adolescents on CT indicates the fracture producing force vector: a biomechanical rationale. *Internet J Spine Surg*. 2005; 1: 1-6.
20) Pedersen AB, Mikkelsen EM, Cronin-Fenton D, et al. Missing data and multiple imputation in clinical epidemiological research. *Clin Epidemiol*. 2017; 9: 157-166.

21) DataCamp.com. R Documentation and manuals. https://www.rdocumentation.org/. Accessed May 26, 2020.

Supplemental Digital Content1. Explanations for the packages and function in R[21)]

"mice" : mice is a package that calculates multivariate imputation by chained equations. The package was used to impute missing data including continuous, binary, and ordered categorical data with stochastic regression imputation.

"rpart" : rpart is a package for recursive partitioning and regression trees. It was employed to conduct CART analysis and build regression models drawn as binary trees.

"ROCR" : ROCR is used for visualizing the performance of scoring classifiers. This package was used to calculate sensitivity, specificity, accuracy, and the area under the ROC curve.

"epiR" : epiR is a package of tools for epidemiological data analysis. This package was used to calculate prevalence ratios from contingency tables.

"sample" : sample is used to return random samples and permutations. This function was used to randomly select a sample of specified size without replacement.