Osteoporosis and Sarcopenia 10 (2024) 40–44

Contents lists available at ScienceDirect

Osteoporosis and Sarcopenia

journal homepage: www.elsevier.com/locate/afos

Original article

Temporal validation of a clinical prediction rule for distinguishing locomotive syndromes in community-dwelling older adults: A cross-sectional study from the DETECt-L study

Shigeharu Tanaka [a,b], Ryo Tanaka [b,*], Hungu Jung [b], Shunsuke Yamashina [b], Yu Inoue [c], Kazuhiko Hirata [d], Kai Ushio [e], Yasunari Ikuta [f], Yukio Mikami [e], Nobuo Adachi [f]

[a] Physical Therapy Major, School of Rehabilitation, Kanagawa University of Human Services, Yokosuka, Kanagawa, Japan
[b] Graduate School of Humanities and Social Sciences, Hiroshima University, Higashi Hiroshima, Hiroshima, Japan
[c] Department of Physical Therapy, School of Health Science and Social Welfare, Kibi International University, Takahashi, Okayama, Japan
[d] Division of Rehabilitation, Department of Clinical Practice and Support, Hiroshima University Hospital, Hiroshima, Hiroshima, Japan
[e] Department of Rehabilitation Medicine, Hiroshima University Hospital, Hiroshima, Hiroshima, Japan
[f] Department of Orthopaedic Surgery, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Hiroshima, Japan

ARTICLE INFO

Keywords:
Locomotive syndrome
Clinical prediction rule
Validation
Aging

ABSTRACT

Objectives: Clinical prediction rules are used to discriminate patients with locomotive syndrome and may enable early detection. This study aimed to validate the clinical predictive rules for locomotive syndrome in community-dwelling older adults.

Methods: We assessed the clinical prediction rules for locomotive syndrome in a cross-sectional setting. The age, sex, and body mass index of participants were recorded. Five physical function tests–grip strength, single-leg standing time, timed up-and-go test, and preferred and maximum walking speeds–were measured as predictive factors. Three previously developed clinical prediction models for determining the severity of locomotive syndrome were assessed using a decision tree analysis. To assess validity, the sensitivity, specificity, likelihood ratio, and post-test probability of the clinical prediction rules were calculated using receiver operating characteristic curve analysis for each model.

Results: Overall, 280 older adults were included (240 women; mean age, 74.8 ± 5.2 years), and 232 (82.9%), 68 (24.3%), and 28 (10.0%) participants had locomotive syndrome stages ≥ 1, ≥ 2, and = 3, respectively. The areas under the receiver operating characteristics curves were 0.701, 0.709, and 0.603, in models 1, 2, and 3, respectively. The accuracies of models 1 and 2 were moderate.

Conclusions: These findings indicate that the models are reliable for community-dwelling older adults.

1. Introduction

Aging population is a social problem in industrialized countries. The National Institute of Population and Social Security Research reported that the aging rate in Japan was approximately 28% in 2017 and will reach 35% in 2040 [1]; this means that Japan has one of the highest aging rates globally. In response, the Japanese Orthopaedic Association (JOA) proposed "locomotive syndrome (LS)" to develop simple pretests to assess individual locomotive abilities and determine which individuals would need nursing care [2]. The number of older adults with LS is considered high; a previous longitudinal study showed that 69.8% of community-dwelling older adults have LS [3]. Moreover, previous research studying the association between LS and health issues in older adults revealed that LS is linked to chronic pain, osteoporosis, lumbar diseases, an increased risk of falling, and a poor quality of life [4,5]. Early detection and accurate assessment of LS by healthcare providers may be crucial to prevent health problems in older adults.

Physical function assessments, such as grip strength, single-leg standing (SLS) time, timed up-and-go test (TUG), and preferred and maximum walking speed, were used as predictive factors for LS. These tests are simple, inexpensive, and have a low physical burden on older adults; therefore, they are often used in clinical practice. A previous

Peer review under responsibility of The Korean Society of Osteoporosis.
* Corresponding author. Graduate School of Humanities and Social Sciences, Hiroshima University, Higashi Hiroshima, Hiroshima, Japan.
E-mail address: ryotana@hiroshima-u.ac.jp (R. Tanaka).

https://doi.org/10.1016/j.afos.2024.02.003
Received 9 December 2023; Received in revised form 29 January 2024; Accepted 27 February 2024
Available online 15 March 2024

study examined a clinical prediction rule (CPR) to distinguish LS in 186 community-dwelling older adults and confirmed that SLS time, grip strength, and TUG were significant predictors of LS [6]. In addition, the accuracy of the prediction model in a previous study was moderate. Therefore, CPR may help diagnose LS in older adults and may be beneficial to healthcare providers in developing appropriate interventions in clinical practice.

CPRs are generally used to predict the absolute risk of clinical conditions and future health outcomes. The development of a CPR comprises three main stages: derivation, validation, and impact analysis [7]. To maximize their clinical usefulness, CPRs must be rigorously developed and validated, and their impact on clinical practice and patient outcomes must be evaluated. In this modeling strategy, a CPR for LS has already been derived in a previous study [6]; therefore, the validation and impact analysis steps are required before this CPR can be recommended for use in clinical practice.

Generally, there are three designs for examining CPR validation after derivation: temporal, geographic, and domain [8]. These three designs can be used to evaluate the generalizability or transportability of CPR model performance in patients from different but plausibly related settings. Temporal validation is a common design approach, with only a difference in calendar time between the derivation and validation datasets. Temporal validation involves assessing the reproducibility of a rule in different populations, and this design ensures evidence for clinical practice [9]. To date, CPR for LS has not been validated. This study aimed to validate the use of CPR in determining LS in community-dwelling older adults.

2. Methods

2.1. Ethics

All measured parameters were essential for assessing functional status and were not harmful. All the procedures were performed in accordance with the Declaration of Helsinki (1975). Written informed consent was obtained from all the participants. This study was approved by the ethical review board of this institution (E2022-0086-04). All patients agreed to participate in the study and provided written informed consent.

2.2. Study setting

This was a cross-sectional validation study of CPR for determining LS. The data were extracted from the Study for Diagnosis, Early detection, and Treatment of locomotive syndrome using the Epidemiological Cohort "DETECt-L″ database, and participants were recruited using an explanatory poster between May 2021 and December 2022. Staff at community centers and gymnasiums throughout Hiroshima City, Japan as well as exercise instructors who sponsor exercise classes in the community helped recruit and measure the variables. The participants were free of charge.

2.3. Participants

The inclusion criteria for this study were community-dwelling persons $\geq$ 65 years old and those with independent mobility. Participants with suspected cognitive impairment or serious illnesses (unstable cardiovascular disease, stroke, severe respiratory impairment, Parkinson's disease, diabetic peripheral neuropathy, or rheumatoid arthritis) were excluded. This was because cognitive impairment could have resulted in unreliable or invalid responses to the questionnaire, whereas serious diseases could have led to falls or disease exacerbation during assessment.

To examine temporal validation, the criteria for participants were set according to a previous study [6]. Both this and our previous studies developed the CPRs in the region from which the participants were recruited, with similar recruitment methods as well as the selection and exclusion criteria for the participants. However, the timing of the assessments was different for examining temporal validation. Another participant who had not participated in previous studies was included.

2.4. Assessments

Participants' age, sex, and body mass index (BMI) were recorded. Five physical function tests—SLS time, grip strength, preferred and maximum walking speeds, and TUG test—were evaluated as predictive factors.

SLS time was measured twice, and the larger value of the two measurements was used in the statistical analyses. After determining the supporting leg, the participants assumed a single-leg standing posture when instructed by the evaluator to raise one leg. The raised leg was maintained at five cm above the ground. The time required to maintain this position was measured for a maximum of 60 s. Test trials with fractions from the second round were rounded down. The SLS was a valid and reliable test in older adults in previous studies [10,11].

Grip strength was measured using a grip strength meter (TKK 5401 Grip-D; Takei, Niigata, Japan). Strength in kilograms was measured for each hand, and the average of the two measurements was used for subsequent analysis. The participants grasped the grip strength meter with its pointer on the outside. The interphalangeal joints of the fingers were adjusted such that they were almost at right angles before the measurements. A previous study demonstrated the high reliability and validity of the grip strength measurement method [12].

Preferred and maximum walking speeds were measured. Each test was measured once, and the time was recorded in seconds. The participants were instructed to walk along a 5-m pathway. Walking time is an easy test, and its reliability and validity have been confirmed [13].

The TUG test was performed twice, and the smaller value was used for subsequent analyses. The TUG test involves rising from a 40 cm-high chair and walking 3 m, turning around, walking back to the chair, and sitting down. The TUG test was performed using a stopwatch and recorded as the time(s) required to complete the task. The participants were given the following standard instructions: "Walk to the line, turn around, return, and then sit back in a chair. Walk as fast as possible." Previous studies have demonstrated excellent test-retest reliability and good validity of the TUG test in older adults [13,14].

The stand-up test, 2-step test, and 25-item Geriatric Locomotive Function Scale were used to determine LS, as previously described in a study that derived the CPR for LS [6]. These 3 tests are valid, reliable, and feasible in evaluating the risk of LS [15,16]. LS severity was staged from 0 (non-LS) to 3 (most severe) based on these tests. The most severe stages were used in the subsequent analyses.

In a previous study [6], 3 models were built based on the LS severity: Model 1 was used to identify participants with and without LS, using a binary dependent variable with LS severity of either 0 or $\geq$ 1; Model 2 was used to categorize moderate or high LS severity, with a binary dependent variable of LS severity $\leq$ 1 or $\geq$ 2; finally, Model 3 used a binary dependent variable with LS severity of $\leq$ 2 or 3. This study used the same model classification methods. LS severity was evaluated using three tests, including the stand-up test, 2-step test, and 25-item Geriatric Locomotive Function Scale, which required instruments, questionnaires, and time for measurement. The LS models helped improve the accuracy and speed of LS screening in clinical practice. In particular, they are useful when performing mass examinations on several participants, such as group assessments of older adults.

2.5. CPR for distinguishing the LS condition

A previous study [6] developed three CPR models for each LS severity using decision tree analysis; these models are referred to as L-TreeS in this study. The L-treeS identified the following significant variables: SLS time and grip strength in Model 1; SLS time, TUG, and

Table 1
Cut-off values of the L-treeS.

Model	Point		Optimal cut-off point
Model 1			> 1 point
	1 point	SLS time > 59.4 s and grip strength > 37.8 kg	
	2 points	SLS time > 59.4 s and grip strength ≤ 37.8 kg	
	3 points	SLS ≤ 59.4 s	
Model 2			> 3 points
	1 point	SLS time > 12.6 s and SLS time > 55.3 s	
	2 points	SLS time > 12.6 s and SLS time ≤ 55.3 s	
	3 points	SLS time ≤ 12.6 s and TUG ≤ 7.9 s	
	4 points	SLS time ≤ 12.6 s and TUG > 7.9 s	
Model 3			> 2 points
	1 point	Predictive value in Model 2 = 0	
	2 points	Predictive value in Model 2 = 1 and maximum walking speed ≤ 3.75 s	
	3 points	Predictive value in Model 2 = 1 and maximum walking speed > 3.75 s	

SLS, single-leg standing.

predictive value in Model 2; and maximum walking time in Model 3.

In this temporal validation study, the validity of CPRs was examined by distributing scores in descending order, starting from the terminal node with the largest percentage of positive or more severe LS cases (eg, in model 1, participants with SLS ≤ 59.4 s were scored 3 points, participants with SLS time > 59.4 s and grip strength ≤ 37.8 kg were scored 2 points, and participants with SLS time > 59.4 s and grip strength > 37.8 kg were scored 1 point) (Table 1). The participants were scored based on the CPR criteria for LS.

2.6. Sample size

A previous cohort study found that 69.8% of community-dwelling older adults in Japan had LS [3]. Therefore, we assumed that the ratio of older adults with LS to those without LS would be 2.5:1. The alpha value and power were set to 0.05 and 0.80, respectively. The null hypothesis value for the area under the receiver operating characteristic (AUROC) was set at 0.70. The target sample size was 84. Therefore, additional participants were recruited to ensure an adequate sample size.

Table 2
Descriptive statistics of all participants.

	All (N = 280)	Women (N = 240)	Men (N = 40)
Age, yrs	74.8 (5.2)	74.6 (5.3)	75.9 (3.8)
SLS time, s	32.5 (22.3)	33.7 (22.3)	25.1 (21.0)
Grip strength, kg	23.5 (6.0)	21.7 (4.0)	34.1 (5.1)
Walking time (preferred), s	3.8 (0.8)	3.8 (0.7)	3.8 (0.9)
Walking time (max), s	2.9 (0.6)	2.9 (0.6)	2.8 (0.7)
TUG, s	6.4 (1.4)	6.5 (1.3)	6.1 (2.0)
Severity of locomotive syndrome, N (%)			
≥ Stage 1	232 (82.9)	198 (82.5)	34 (85.0)
≥ Stage 2	68 (24.3)	61 (25.4)	7 (17.5)
= Stage 3	28 (10.0)	25 (10.4)	3 (7.5)

SLS, single-leg standing; TUG, timed up-and-go test.

2.7. Statistical analysis

Receiver operating characteristic (ROC) curve analysis was used to determine the sensitivity, specificity, likelihood ratio (LR), and post-test probability of the CPRs for each model. The cutoff values were set based on the results of L-treeS [6]. The optimal cut-off points were > 1 point in Model 1, > 3 points in Model 2, and > 2 points in Model 3 (Table 1). The area under the ROC curve (AUROC) with a 95% confidence interval (CI) was used to assess the accuracy of the models. The AUROC could be distinguished from 0 to 1 and was categorized as low (0.5 < AUROC < 0.7), moderate (0.7 < AUROC < 0.9), and high (0.9 < AUROC < 1) [17]. Statistical analyses were performed using SPSS version 25 software for Windows (IBM Corp., Armonk, NY, USA), and statistical significance was set at $P < 0.05$.

3. Results

In total, 280 older adults (240 women, 40 men; mean age, 74.8 ± 5.2 years) participated in this study. Descriptive statistics are shown in Table 2. Overall, 232 (82.9%), 68 (24.3%), and 28 (10.0%) participants had LS stages ≥ 1, ≥ 2, and = 3, respectively.

The AUROC values were 0.701 (95% CI, 0.614–0.788; P < 0.01), 0.709 (95% CI, 0.635–0.783; P < 0.01), and 0.603 (95% CI, 0.480–0.727; P = 0.07), in Models 1, 2, and 3, respectively (Table 3). The accuracies of Models 1 and 2 were determined to be moderate based on the AUROC results. The AUROC value in Model 3 was not significant. The post-test probability test revealed positive predictive values (PPV) of 83.5%, 78.9%, and 55.6% and negative predictive values (NPV) of 100%, 79.7%, and 91.5% for Models 1, 2, and 3, respectively.

4. Discussion

This study examined the temporal validity of previously derived CPRs for identifying LS in community-dwelling older adults. A previous study, which included 186 participants, presented CPRs for each LS severity, and the AUROC values were 0.737, 0.763, and 0.704, respectively, for Models 1, 2, and 3 [6]. The present study assessed the validity

Table 3
Diagnostic characteristics of each model.

								Post-test Probability	
Models	AUROC	95% CI	P-value	Sensitivity	Specificity	LR+	LR-	PPV, %	NPV, %
Model 1	0.701	0.614–0.788	< 0.01	1.000	0.042	1.043	0.000	0.835	1.000
Model 2	0.709	0.635–0.783	< 0.01	0.221	0.981	11.691	0.794	0.789	0.797
Model 3	0.603	0.480–0.727	P = 0.07	0.179	0.984	11.250	0.835	0.556	0.915

AUROC, area under the receiver operating characteristic; 95% CI, 95% confidence interval; LR, likelihood ratio; PPV, positive predictive value; NPV, negative predictive value.

of the same models in 280 community-dwelling older adults, with AUROC values of 0.701, 0.709, and 0.603 for Models 1, 2, and 3, respectively. This study confirms that Models 1 and 2 have moderate validity. This study confirmed the temporal validity of CPR for LS and demonstrated the reliability of these models in clinical practice. To the best of our knowledge, the literature reported no validated CPR for LS. This study is the first to validate CPR for LS, and our results will help healthcare providers in managing and supporting older adults with health problems.

Temporal validation used only a difference in the assessment timing between the derivation and validation datasets. This study included 280 participants (mean age, 74.8 ± 5.2 years) and 232 (82.9%), 68 (24.3%), and 28 (10.0%) participants had LS stages ≥ 1, ≥ 2, and = 3, respectively. A previous study derived a CPR for LS included 186 participants (mean age, 75.0 ± 6.5 years) and 150 (80.6%), 44 (23.7%), and 15 (8.1%) participants had LS stages ≥ 1, ≥ 2, and = 3, respectively; this meant participants in the present study were similar to those in a previous study. Thus, the participants in this study were considered suitable for temporal validation, and the results were highly reliable.

In this study, we examined the temporal validity of L-treeS, which was derived using decision tree analysis in a previous study. In general, machine learning including decision tree analysis is prone to overfitting [18]. Overfitted CPR models may lead to inadequate conclusions, which may incorrectly or even harmfully shape clinical decision-making. In this temporal validation study, the same levels of AUROC could be confirmed using different participants, indicating a low level of overfitting in the derived CPR for the LS. Addressing overfitting concerns will lead to recommendations for ensuring the usefulness of CPR models for LS in clinical practice.

The pre-test probability of participants with LS was 82.9%, the PPV in Model 1 was 83.5%, similar to the pre-test probability. These results indicated that Model 1 was inappropriate for identifying older adults with LS. However, the NPV was 100% for Model 1. Therefore, Model 1 was useful for distinguishing older adults without LS. A previous study by Kobayashi et al. [19] in 2022 indicated that SLS time was a significant predictive factor for identifying LS, with an AUROC of 0.66. In this study, the AUROC for Model 1 was 0.701. This validation study provides additional evidence for a more accurate distinction between LS severities.

Moreover, the pre-test probability of participants with LS stage ≥ 2 was 24.3% in this study, and Model 2 has a PPV of 78.9%. Model 2 set as the dependent variable binary LS severity of ≤ 1 or ≥ 2; this model was used to identify whether the severity of LS was moderate or high. LS stage 2 is treated as progressive LS with possible musculoskeletal degeneration that requires clinical consultation [20]. Therefore, Model 2 is useful for identifying older adults with advanced LS and will aid healthcare providers in selecting appropriate treatment plans in clinical practice.

This study validated the L-treeS by combining two tests: SLS time and grip strength in Model 1 and SLS time and TUG in Model 2, which provided statistical accuracy. In clinical practice, SLS time should first be selected to distinguish LS, as SLS time was a significant predictive factor in both Models 1 and 2. If older adults show an SLS time of > 59.5 s, healthcare providers can include a grip strength test for greater diagnostic accuracy. However, if older adults show an SLS time ≤ 12.6 s, healthcare providers can include TUG to determine whether the individual has progressive LS. A previous study indicated that SLS time was a significant predictive factor for LS [21], and the present study showed that these simple combinations of tests were useful for detecting the severity of LS with moderate predictive accuracy, validating the LS models. These findings contribute to the evidence-based care of older adults with LS in clinical practice.

To maximize the potential clinical usefulness of CPR, impact analysis should be performed. Impact analyses are the most efficient methods for assessing whether incorporating CPRs into the decision-making process improves clinical practice. Although there is no perfectly established methodology for the study design of impact analysis, Wallace et al. [22] provided a framework for impact analysis. In addition, a previous study showed that locomotive training (LOCO-Tre), consisting of squats and single-leg standing exercises, is an effective intervention for older adults with LS [4]. Future studies should evaluate whether the use of L-treeS in clinical practice facilitates the transition to more effective interventions for LS.

This study has several limitations. First, although the AUROCs in Models 1 and 2 were significant, those in Model 3 were not. One reason that the AUROC value of Model 3 was not significant is the insufficient number of participants in it. Model 3 investigated a binary LS severity of ≤ 2 or 3. Thus, accurately predicting the most severe LS was limited. Health providers need to pay attention to this point in clinical practice. Additionally, this study only examined temporal validation, although there have been some study designs for validation. Therefore, further validation studies, including those for Model 3, are required. Secondly, the proportion of men in this study was 14.3%. This may have influenced the distinction of LS among male participants. Third, although this study focused on older adults, approximately 20% of young and middle-aged individuals have LS [23]. Therefore, to develop an appropriate prevention program, L-treeS should be validated for domain validation, which means validation in other populations (eg, young and middle-aged populations) [8]. Fourth, models were not created separately for males and females. Grip strength differs between males and females, thus modeling them separately may be more accurate. Further studies are required to overcome these limitations.

5. Conclusions

This study examined the temporal validity of the CPRs for identifying LS in community-dwelling older adults. This study revealed that Models 1 and 2 have moderate validity. This study is the first to validate CPR for LS, and our results will help healthcare providers in managing and supporting older adults with health problems.

CRediT author statement

Ryo Tanaka: Conceptualization, Methodology, Investigation, Writing - Review & Editing, Project administration, Funding acquisition. **Shigeharu Tanaka**: Conceptualization, Methodology, Writing-original draft. **Hungu Jung**: Conceptualization, Methodology, Investigation, Writing - Review & Editing. **Shunsuke Yamashina**: Conceptualization, Methodology, Writing - Review & Editing. **Yu Inoue**: Conceptualization, Methodology, Writing - Review & Editing. **Kazuhiro Hirata**: Conceptualization, resources, Writing - Review & Editing. **Kai Ushio**: Conceptualization, Writing - Review & Editing. **Yasunari Ikuta**: Conceptualization, Writing - Review & Editing. **Yukio Mikami**: Conceptualization, Writing - Review & Editing, resources, Supervision. **Nobuo Adachi**: Conceptualization, Writing - Review & Editing, Supervision.

Conflicts of interest

The authors declare no competing interests.

Acknowledgments

The authors thank Mr. Kazuaki Hamada, Mr. Kenta Hirohama, and Ms. Sayumi Amagata for their assistance with data collection. This work was supported by the MHLW FA program (grant number 22FA1003) and the Chugoku Regional Innovation Research Center (Hiroshima, Japan). The funders had no role in the study design, data collection, statistical analysis, decision to publish, or manuscript preparation. **ORCID** Ryo Tanaka: 0000-0002-1349-3102. Shigeharu Tanaka: 0000-0003-4397-0130. Hungu Jung: 0000-0001-7500-7029. Shunsuke Yamashina: 0000-0003-0560-0702. Yu Inoue: 0000-0001-9920-9442. Kazuhiko

Hirata: 0009-0007-1794-3856. Kai Ushio: 0000-0001-5195-1808. Yasunari Ikuta: 0000-0002-8419-9154. Yukio Mikami: 0000-0001-6855-8439. Nobuo Adachi: 0000-0002-5570-3729.

References

[1] National Institute of Population and Social Security Research [Internet]. Population projections for Japan: 2016–2065. Population Research Series No. 336. [cited 2023 Feb 3]. Available from: http://www.ipss.go.jp/.
[2] Nakamura KA. A "super-aged" society and the "locomotive syndrome". J Orthop Sci 2008;13:1–2.
[3] Yoshimura N, Muraki S, Nakamura K, Tanaka S. Epidemiology of the locomotive syndrome: the research on osteoarthritis/osteoporosis against disability study 2005–2015. Mod Rheumatol 2017;27:1–7.
[4] Matsumoto H, Hagino H, Wada T, Kobayashi E. Locomotive syndrome presents a risk for falls and fractures in the elderly Japanese population. Osteoporos Sarcopenia 2016;2:156–63.
[5] Arbex MCFB, Okazaki JEF, Tavares DRB, Figueiredo Bersani AL, Santos FC. Locomotive syndrome is associated with chronic pain and poor quality of life in Brazilian oldest old: LOCOMOV Project. J Orthop Sci 2021;26:162–6.
[6] Tanaka S, Jung H, Yamashina S, Inoue Y, Nakashima N, Tanaka R. Clinical prediction rule for locomotive syndrome in older adults: a decision tree model. J Orthop Sci 2023;28:886–94.
[7] Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagn Progn Res 2019:3–16.
[8] Binuya MA, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. BMC Med Res Methodol 2022;22:316.
[9] Ramspek CL, Jager KJ, Dekker FW, Zoccali C, Diepen MV. External validation of prognostic models: what, why, how, when and where? Clin Kidney J 2020;14:49–58.
[10] Goldberg A, Casby A, Wasielewski M. Minimum detectable change for single-leg-stance-time in older adults. Gait Posture 2011;33:737–9.
[11] Macrae PG, Lacourse M, Moldavon R. Physical performance measures that predict faller status in community-dwelling older adults. J Orthop Sports Phys Ther 1992;16:123–8.
[12] Abizanda P, Navarro JL, García-Tomás MI, López-Jiménez E, Martínez-Sánchez E, Paterna G. Validity and usefulness of hand-held dynamometry for measuring muscle strength in community-dwelling older persons. Arch Gerontol Geriatr 2012;54:21–7.
[13] Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: six-minute walk test, berg balance Scale, timed up & go test, and gait speeds. Phys Ther 2002;82:128–37.
[14] Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142–8.
[15] Seichi A, Hoshino Y, Doi T, Akai M, Tobimatsu Y, Iwaya T. Development of a screening tool for risk of locomotive syndrome in the elderly: the 25-question Geriatric Locomotive Function Scale. J Orthop Sci 2012;17:163–72.
[16] Ogata T, Muranaga S, Ishibashi H, Ohe T, Izumida R, Yoshimura N, et al. Development of a screening program to assess motor function in the adult population: a cross-sectional observational study. J Orthop Sci 2015;20:888–95.
[17] Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatr 2007;96:644–7.
[18] Kernbach JM, Staartjes VE. Foundations of machine learning-based clinical prediction modeling: Part II-Generalization and Overfitting. Acta Neurochir Suppl 2022;134:15–21.
[19] Kobayashi T, Takenaka H, Tatematsu N, Sugiura H, Noguchi T, Nojima I. Association between locomotive syndrome stage 1 and physical performance measures among the older people living in the community. Rigakuryohogaku 2022;49:212–9 (in Japanese).
[20] Ishibashi H. Locomotive syndrome in Japan. Osteoporos Sarcopenia 2018;4:86–94.
[21] Kobayashi K, Ando K, Tsushima M, Machino M, Ota K, Morozumi M, et al. Predictors of locomotive syndrome in community-living people: a prospective five-year longitudinal study. Mod Rheumatol 2019;29:669–75.
[22] Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, et al. Framework for the impact analysis and implementation of clinical prediction rules (CPRs). BMC Med Inf Decis Making 2011;11:62.
[23] Nishimura A, Ohtsuki M, Kato T, Nagao R, Ito N, Kato K, et al. Locomotive syndrome testing in young and middle adulthood. Mod Rheumatol 2020;30:178–83.