Gait & Posture 68 (2019) 335–339

Contents lists available at ScienceDirect

Gait & Posture

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

Full length article

Concurrent validity and measurement error of stair climb test in people with pre-radiographic to mild knee osteoarthritis

Hirotaka Iijima[a,b,c,*], Kanako Shimoura[b], Ryo Eguchi[d], Tomoki Aoyama[b], Masaki Takahashi[a]

[a] Department of System Design Engineering, Faculty of Science and Technology, Keio University, Yokohama, Japan
[b] Department of Physical Therapy, Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan
[c] Japan Society for the Promotion of Science, Tokyo, Japan
[d] School of Science for Open and Environmental Systems, Graduate School of Science and Technology, Keio University, Yokohama, Japan

ARTICLE INFO

Keywords:
Knee osteoarthritis
Stair climb test
Validity
Bland-Altman plot

ABSTRACT

Background: Stair climbing is the task first affected in patients with knee osteoarthritis (OA); therefore, the precise measurement of time required to climb stairs is important to identify mobility limitations, particularly in the early phase of knee OA.
Research question: This study aimed to examine the test-retest reliability, measurement error, and concurrent validity of the stopwatch-based stair-climb test (SCT) in adults with pre-radiographic to mild knee OA.
Methods: Fifty-nine participants (mean age, 59.1 [range, 50–69] years; 72.9% female) with Kellgren and Lawrence grade ≤2 disease underwent an 11-step SCT (11-SCT) in accordance with the Osteoarthritis Research Society International recommended method while wearing pressure sensor-mounted standard shoes that is used as a gold standard procedure. Test-retest reliability, measurement errors, and the concurrent validity of the stopwatch-based 11-SCT were evaluated.
Results: The test-retest reliability of the stopwatch-based 11-SCT was excellent (intra-class correlation coefficient$_{1,1}$ [ICC$_{1,1}$], 0.952; 95% confidence interval [CI], 0.560 to 0.985; $p < 0.001$) and the minimal detectable change$_{95}$ was 0.102 s. Concurrent validity was excellent (ICC$_{2,1}$: 0.957; 95% CI: 0.661 to 0.986; $p < 0.001$).
Significance: The stopwatch-based 11-SCT had high test-retest reliability and high concurrent validity, which justify its clinical use for identifying mobility limitations in individuals with pre-radiographic to mild knee OA. A difference of 0.2 s in the stopwatch-based 11-SCT time would be considered a true difference beyond a 95% measurement error.

1. Introduction

Knee osteoarthritis (OA) is the leading cause of knee pain and mobility limitations [1]. The measurement of treatment outcomes and mobility limitations over time is a critical component of clinical practice and research in patients with knee OA. Stair climbing is a common and frequent activity that is the first task affected in patients with knee OA [2]. A recently published meta-analysis revealed that a poor stair climbing capacity was associated with established knee OA [3]. Considering that stair climbing is biomechanically and physiologically more challenging than level walking [4,5], a reliable and valid measurement of one's stair climbing capacity is important in evaluating knee OA-related mobility limitations and rehabilitation efficacy in clinical practice and research and may be more useful than analysis of other activities, such as level walking.

Based on currently available evidence and expert consensus, the stopwatch-based stair-climb test (SCT) is recommended as a performance-based test to assess mobility limitations in patients with knee OA [6]. Previous studies reported excellent test-retest reliability and/or measurement errors using the stopwatch-based SCT in patients with hip or knee OA [7], those with severe knee OA scheduled for total knee arthroplasty (TKA) [8], and those with severe OA scheduled for TKA or total hip arthroplasty (THA) [9]. These studies ensured the reliability of the SCT for identigying mobility limitation in these patients. However, no study to date has assessed these measurements in patients with the earlier stage of knee OA (i.e., pre-radiographic and mild knee OA), as currently recommended OA prevention studies [10]. Radiographic disease severity would be expected to influence SCT time variability; thus, the determination of reliability in patients with earlier-stage disease is required. Furthermore, we were unable to identify any studies

* Corresponding author at: 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan.
E-mail addresses: iijima.hirotaka.4m@yt.sd.keio.ac.jp (H. Iijima), shimoura.kanako.53s@st.kyoto-u.ac.jp (K. Shimoura), eguchi.ryo@keio.jp (R. Eguchi), aoyama.tomoki.4e@kyoto-u.ac.jp (T. Aoyama), takahashi@sd.keio.ac.jp (M. Takahashi).

https://doi.org/10.1016/j.gaitpost.2018.12.014
Received 22 July 2018; Received in revised form 9 November 2018; Accepted 8 December 2018

that compared the simple clinical measurement of SCT to a "gold standard" procedure. While reliability refers to the *consistency* or *stability* of measurement, validity refers to the *suitability* or *meaningfulness* of the measurement. Since SCT in clinical practice is evaluated using a manual stopwatch, which includes operability error [11], investigating the concurrent validity (i.e., the manually recorded SCT time compared to the gold standard procedure) further ensures its clinical use for identifying mobility limitations. Given that patients with less severe radiographic disease respond better to therapeutic interventions than patients with severe disease [12], investigating the test-retest reliability, measurement error, and concurrent validity of SCT and identifying mobility limitations at an earlier stage of knee OA has a therapeutic advantage.

This study aimed to examine the test-retest reliability, measurement error, and concurrent validity of the stopwatch-based SCT in patients with pre-radiographic to mild knee OA. To determine concurrent validity, the stopwatch-based SCT time was compared with measurements made using a pressure sensor-mounted insole (i.e., the gold standard procedure). We hypothesized that stopwatch-based SCT had excellent test-retest reliability and concurrent validity.

2. Methods

2.1. Participants

This study is a secondary cross-sectional analysis of baseline data from a randomized controlled trial (RCT) examining the immediate effects of transcutaneous electrical nerve stimulation (TENS) on knee pain and physical function in individuals who report knee pain [13]. The ethics committee of our university approved the study (approval number: C1349), and written informed consent was obtained from all participants before enrollment. The required sample size for the RCT was calculated based on the data from a previous clinical trial to detect the therapeutic effects of TENS. Details of the sample size calculation were provided in the original RCT [13]. Community dwelling adults reporting knee pain within the prior month were identified through a website. Those interested in this study answered a questionnaire concerning physical condition, and those who met the inclusion criteria were invited to the university in February 2018 to participate in the study. The eligibility criteria included (1) age $\geq$ 50 years; (2) Kellgren and Lawrence (K&L) grade $\leq$2 in one or both knees evaluated using weight-bearing anteroposterior radiographs, and (3) average pain score $\geq$4 but $\leq$9 on a numeric rating scale (0–10 total points) in the prior month. Participants were excluded if they (1) had a history of knee surgery or intra-articular corticosteroid or hyaluronic acid injection within 6 months, (2) had a history of knee joint replacement or tibial osteotomy, (3) were receiving physical therapy, (4) had any other major joint pain (e.g., back, hip, or ankle) with a greater effect than their knee pain on current limits of ability, (5) had severe medical or neurologic conditions (e.g., chronic obstructive pulmonary disease, cardiovascular disease, arteriosclerosis obliterans, cerebrovascular accident, lumbar disc herniation, or rheumatoid arthritis), (6) did not usually use stairs in daily living, and (7) did not have the ability to walk or climb stairs without an ambulatory assistive device and handrail. From the website survey, 59 participants (mean age, 59.1 [range, 50–69] years; 72.9% female) were included in the original RCT [13]. The participants underwent a 1-month observation period, during which they did not receive any treatment for knee pain as part of the study. This was to ensure that only participants with ongoing chronic symptoms were enrolled.

2.2. Study procedure

All tests were administered by trained physical therapists and performed in the same testing session. During the testing session, the subjects first completed a self-reported questionnaire of personal characteristics, the OA-related health domain measure (Japanese Knee Osteoarthritis Measure [JKOM]), and radiographic assessment of disease severity, followed by the SCT. The JKOM is a patient-based self-answered scoring system that assesses "pain and stiffness" (8 questions, 0–32 points), "activities of daily living" (10 questions, 0–40 points), "participation in social activities" (5 questions, 0–20 points), and "general health conditions" (2 questions, 0–8 points), with a maximum score of 100 points, in a person-specific assessment [14]. Radiographic OA severity in both knees was assessed using the anteroposterior short view in the weight-bearing position and the original version of the K&L grading system [15] as described in the previous study [16]. OA severity in the tibiofemoral joint was assessed by an experienced examiner. We previously reported excellent intra-rater (kappa: 0.876; 95% confidence interval [CI]: 0.829 to 0.924) and inter-rater (kappa: 0.845; 95% CI: 0.793 to 0.897) reliability scores [17].

2.3. Instrumentation

Footwear influences balance performance [18,19] and walking speed [19]; thus, to prevent potential bias from the their own footwear, all participants wore pressure sensor (Force register sensors Model 402; Interlink Electronics Inc., CA, USA)–mounted standard shoes (LD AROUND M, Mizuno, Tokyo, Japan) on both feet [20]. A total of 15 sensors were attached to the shoe's insole to allow the precise detection of the initial contact and foot-off time points during SCT. Concurrent validity of the step detection during locomotion has been tested in the laboratory setting and against gold standard procedures such as the force platform [21,22], three-dimensional motion capture systems [23], foot switch [24,25], and pressure-sensing insole [26]. Among these devices, the latter two methods allow for the examination of SCT time on actual stairs outside the motion laboratory. Furthermore, during stair climbing, forefoot contact without heel contact was expected, and a pilot study of one healthy adult (age: 30 years; height: 1.83 m; body mass: 75.0 kg) indeed confirmed forefoot initial contact during stair descent, which cannot be adequately captured by a single foot switch (see Supplemental Appendix, S1). In-shoe foot pressure sensors can more precisely detect foot contact than a single foot sensor; therefore, they were used as a gold standard procedure in this study. The sensors were connected to a flexible circuit board containing a microcontroller and a micro SD card that were placed on the lower legs [20]. Although the footwear was novel for each participant, all subjects verbally reported that the shoes were comfortable and caused no pain. Data were sampled at 100 Hz.

2.4. Procedure for determining 11-SCT time

Each participant underwent the SCT in accordance with the Osteoarthritis Research Society International recommended method (http://oarsi.org/research/physical-performancemeasures), while wearing pressure sensor–mounted standard shoes. To exclude the influence of participant' clothing on stopwatch operability, the pants were rolled up to the knee level. The 11-SCT using a regular stairwell was selected since it represents the stair types and sizes that subjects likely need to manage during daily activities; moreover, previous studies of reliability also used the 11-SCT [7,27]. Participants descended and ascended a flight of stairs consisting of 11 steps with a height of 17 cm, width of 135 cm, and tread of 29 cm each. Although the standard SCT starts with stair ascent, the 11-SCT in this study started with descent because of environment constraints. A pilot study performed in the same place with the conditions of the validation study confirmed that the 11-SCT time in six healthy adults (age: 24.3 $\pm$ 3.93 years; 1 female) did not significantly differ between an ascending (9.31 $\pm$ 1.47 s) and descending (9.17 $\pm$ 1.45 s) start (mean difference: 0.15 $\pm$ 0.43 s; 95% CI: -0.30 to 0.60; p = 0.439). A trained physical therapist measured the time required to perform the 11-SCT in a test-retest session using a stopwatch (TD-392; TANITA Corp., Tokyo,

Japan). On the trained physical therapist's "start" command, each participant was instructed to walk as quickly and safely as possible without running [6] and was subsequently retested to ensure a stable physical condition during the testing period. The use of any walking aid or handrail was prohibited. Participants performed one practice trial for familiarization and to verify safety. The stopwatch-based 11-SCT time was defined as the time from the "start" command to when the subject reached a point with both feet on the top platform. The stopwatch measurement and the gold standard procedure (i.e., insole measurement) were performed simultaneously. Knee pain during the 11-SCT was evaluated after the retest procedures using a visual analog scale (0–100 mm).

To determine insole-based 11-SCT time, the vertical ground reaction force was calculated, as previously described [20]. The initial contact and foot-off on each side were subsequently defined as the time point at which the vertical ground reaction force exceeded or was below the 1% threshold of estimated maximum ground reaction force using MATLAB version R2017a (MathWorks Inc., Natick, MA, USA). Insole-based 11-SCT time was defined as the time from the initial foot-off on the top platform to the point of foot contact with both feet on the top platform.

2.5. Calculation of test-retest reliability, measurement error, and concurrent validity

Test-retest reliability of the stopwatch-based 11-SCT time was calculated using the intra-class correlation coefficient ($ICC_{1,1}$) with a 95% CI for a one-way random-effects model and absolute agreement. Measurement error in the stopwatch-based 11-SCT time was expressed as the standard error of measurement (SEM). To quantify inherent variability, the minimal detectable changes (MDC_{90} and MDC_{95}) were calculated.

A linear regression analysis was performed to estimate the predictive ability of the stopwatch-based 11-SCT for the insole-based 11-SCT and a regression equation was determined. We checked the regression model features via residuals versus fitted values (i.e., residuals had to be normally distributed around zero), and independence between observations. Concurrent validity (inter-method reliability) was also calculated using $ICC_{2,1}$ with a 95% CI for a two-way random-effects model and absolute agreement. Bland-Altman plots were used to assess agreement between the stopwatch- and insole-based 11-SCT times. The difference between and mean of the two measurement methods were plotted against each other. The average difference and 95% CI were calculated. The proportional bias and radiographic disease severity effect on the difference between the two methods were also evaluated. In test-retest reliability and concurrent validity, the interpretation of ICC was based on inspection of the lower one-side 95% CI set at a minimum acceptable level of 0.700 and point estimates set a sufficient level of 0.800 or more [28,29]. All statistical analyses were performed using JMP Pro 13.0 (SAS Institute, 100 SAS Campus Drive Cary, NC, USA). A p value < 0.05 was considered statistically significant.

3. Results

Fifty-nine participants were enrolled in the original RCT and all participants were included in this cross-sectional secondary analysis. Table 1 shows the participants' characteristics. All patients with K&L grade 1 or 2 disease (61.0%) had medial OA in the index knee.

3.1. Stopwatch-based 11-SCT had high test-retest reliability

The point estimate of test-retest reliability was excellent with excellent agreement [29] ($ICC_{1,1}$: 0.952; 95% CI: 0.560 to 0.985; $p <$ 0.001), although the lower limit of 95% CI did not reach 0.700. The SEM for test-retest reliability was 0.036 s and the resulting MDC_{90} and MDC_{95} values were 0.086 s and 0.102 s, respectively.

Table 1
Participant characteristics (n = 59).

Age, years	59.1 ± 6.05
Female, no. (%)	43 (72.9)
Height, m	1.60 ± 0.08
Mass, kg	58.7 ± 10.5
BMI, kg/m^2	22.7 ± 3.55
Index knee K&L grade, no. (%)	
Grade 0	23 (38.9)
Grade 1	27 (45.8)
Grade 2	9 (15.3)
Bilateral knee OA, no. (%)*	6 (10.2)
VAS score for knee pain during stair climbing, mm	15.7 ± 17.3; 10.0 [0, 67]†
JKOM, points	
Pain and stiffness	7.00 ± 3.89; 7 [0,22]†
Activities of daily living	2.98 ± 3.09; 2 [0,14]†
Participation in social activities	2.44 ± 1.95; 2 [0,9]†
General health conditions	1.97 ± 0.98; 2 [0,4]†
Total score	14.4 ± 7.60; 13 [3, 49]†
11-SCT, s	
Stopwatch-based 11-SCT time	
Test	8.88 ± 1.35
Retest	9.20 ± 1.40
Average	9.04 ± 1.37
Insole-based 11-SCT time	
Test	9.37 ± 1.50
Retest	9.40 ± 1.42
Average	9.38 ± 1.43

BMI, body mass index; JKOM, Japanese Knee Osteoarthritis Measure; K&L grade, Kellgren and Lawrence grade; OA, osteoarthritis; VAS, visual analog scale; 11-SCT, 11-step stair climb test.
Except where otherwise indicated, values are mean ± SD.
* Bilateral knee OA is defined as K&L grade ≥ 2 in both knees.
† Median [lower range, upper range] is also provided because of the scattered distribution of the answered item.

3.2. Stopwatch-based 11-SCT had high validity but systematic bias

Fig. 1 compares the two measurement methods presented as a regression analysis (A) and Bland-Altman plot (B). The point estimate of inter-method reliability was excellent ($ICC_{2,1}$: 0.957; 95% CI: 0.661 to 0.986; $p <$ 0.001), although the lower limit of 95% CI did not reach 0.700. Linear regression revealed that the stopwatch-based 11-SCT had a high predictive ability for insole-based 11-SCT with a high coefficient of determination (R^2: 0.962). The Bland-Altman analysis revealed that the mean stopwatch-based 11-SCT time was 0.304 s (95% CI: 0.230 to 0.379 s) shorter than the mean insole-based 11-SCT time, with limits of agreement of -0.254 to 0.861 s. A significant proportional bias was confirmed (Fig. 1B, right panel). Radiographic disease severity had a small and non-significant effect on the difference in 11-SCT between the two measurement methods.

4. Discussion

This study aimed to test the hypothesis that the stopwatch-based SCT had excellent test-retest reliability and concurrent validity (i.e., compared with a gold standard sensor-mounted insole procedure for determining 11-SCT) and determine the absolute measurement error of the stopwatch-based 11-SCT in 59 subjects with pre-radiographic to mild knee OA. Based on the point estimates of the ICC, stopwatch measurement had excellent reliability and concurrent validity [29], thus supporting the aforementioned hypotheses. MDC_{95} was 0.102 s, indicating that a difference of 0.2 s in the stopwatch-based 11-SCT time would be considered a true difference beyond the 95% measurement error.

4.1. Test-retest reliability and measurement error

The excellent test-retest reliability and measurement error (MDC_{90}: 0.086 s; MDC_{95}: 0.102 s) in this study were in line with or much better

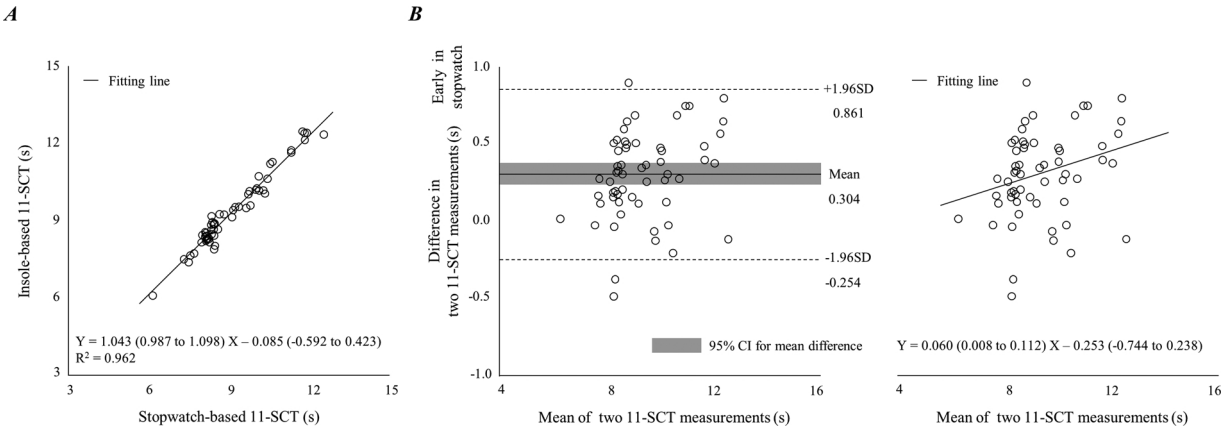

Fig. 1. Comparison of two measurements methods using regression analysis (A) and a Bland-Altman plot (B).
The solid line represents the regression line.
The regression equation is expressed as: Y = a (95% CI) X + b (95% CI).
95% CI, 95% confidence interval; SCT, stair-climb test.

than those reported in previous studies of hip and knee OA (n = 50; MDC_{95}: 2.77 s) [7], severe knee OA scheduled for TKA (n = 30; MDC_{95}: 5.50 s) [8], and severe OA scheduled for primary TKA or THA (n = 150; MDC_{90}: 5.49 s) [9]. In this study, the retest of 11-SCT was performed immediately after the first test to avoid changes in participant physical condition, which might explain the higher test-retest reliability and measurement error than those of previous studies. Furthermore, 11-SCT time and knee pain intensity was much less than that reported in these studies [7,8], which implies relatively better knee pain and function in the participants in this study and might influence main results.

The small measurement error of 11-SCT would enable us to use the test as an adjunct to identify knee OA-related mobility limitations and rehabilitation efficacy. Although there is a lack of adequate pooled data on the SCT in knee OA [3], the difference in the mean 11-SCT between OA patients and healthy adults was 0.80 s in mild knee OA and 3.10 s in moderate knee OA [30]. Since the MDC_{95} in this study was 0.102 s (i.e., a difference of 0.2 s in stopwatch-based 11-SCT time would be considered a true difference), the stopwatch-based 11-SCT has a distinct ability to identify the presence of OA-related mobility limitations in at least patients with mild to moderate knee OA. Although we are not aware of an RCT investigating rehabilitation efficacy based on SCT in pre-radiographic knee OA, 12 weeks of hip abductor muscle strengthening exercise improved the SCT time by 1.0 s in patients with mild to severe knee OA [31]. The measurement error of 0.102 s is adequate for the identification of therapeutic exercise efficacy in these patients.

4.2. Concurrent validity

Accurate stopwatch-based SCT time is required to identify mobility limitation in clinical practice. Although the ideal option would always be to use a force platform or three-dimensional motion capture system [21–23], these devices cannot be used outside the laboratory. Alternatively, this study used a pressure-sensing insole [20] as a gold standard procedure to identify initial contact in both limbs during locomotion [26], which can be used in actual stairs outside the motion laboratory. Wearing pressure sensor-mounted standard shoes could accurately identify initial contact even when forefoot contact was expected and could exclude influence of patient's own shoes on SCT performance [18,19].

Excellent inter-method agreement for 11-SCT time justifies its use for determining OA-related mobility limitations and rehabilitation outcomes in clinical practice and research. Since no data are available on the concurrent validity of stopwatch-based 11-SCT, the findings of this study cannot be compared to those of other studies. Clinicians should interpret our stopwatch-based 11-SCT findings with caution since there was systematic bias (i.e., stopwatch-based 11-SCT time is 0.304 s shorter than the insole-based 11-SCT time, a gold standard method). Potential sources of the differences between the two measurement methods include evaluator operability of the stopwatch and definition of 11-SCT time. The measurement of stopwatch-based 11-SCT began with the "start" command, which is an earlier time point than that using insole-based 11-SCT, which begins at the initial foot-off on the top platform.

4.3. Limitations

First, knee pain intensity was evaluated only after the retest SCT. Thus, we cannot deny the possibility that test performance may increase pain perception, which may influence the retest results. Nevertheless, the global knee pain score was comparable before and after all baseline performance measurements of the original RCT [13], indicating that the influence of altered pain perception on retest performance would be minimal. Second, 11-SCT in this study was performed while wearing standardized shoes that may not be used in clinical practice. The data should be interpreted with caution since most of the clinicians would perform the stopwatch-based 11-SCT with participants wearing their own footwear, which might influence the SCT time. Third, this study did not calculate the required sample size to establish adequate and minimal acceptable reliability for clinical measurement. The lower limit for the 95% CI of test-retest reliability and concurrent validity did not reach the minimum benchmarks of acceptable reliability (i.e., ICC > 0.700), possibly due to the small sample size. Finally, this study did not determine whether the 11-SCT time can be consistently measured by more than one rater (i.e., inter-examiner reliability). The results of this study do not necessarily ensure the reliability of 11-SCT measurement for knee OA by different clinicians in a clinical setting. Since patients with knee OA tend to participate in both inpatient and outpatient physical therapy, determining inter-examiner reliability is warranted as was done in patients with TKA [27].

5. Conclusion

This study examined the test-retest reliability, measurement error, and concurrent validity of stopwatch-based SCT in patients with pre-radiographic to mild knee OA. Given that the measurement of mobility limitations over time is a critical component of clinical practice and research, high test-retest reliability and concurrent validity of the stopwatch-based 11-SCT justifies its clinical use for identifying mobility limitation in patients with pre-radiographic to mild knee OA. A

difference of 0.2 s in stopwatch-based 11-SCT time would be considered a true difference beyond the 95% measurement error.

Conflict of interest

The authors declare no financial support or other benefits from commercial sources for the work reported in the manuscript or any other financial interests that could create a potential conflict of interest regarding this work.

Author contributions

All authors made substantial contributions to: (1) study conception and design, data acquisition, or data analysis and interpretation; (2) drafting the article or revising it critically for important intellectual content; and (3) providing final approval of the manuscript for submission.

The specific contributions of the authors are as follows:

(1) Study conception and design: HI and KS;
(2) Data analysis and interpretation: HI, KS, RE, TA, and MT;
(3) Article drafting: HI, KS, RE, TA, and MT;
(4) Critical revision of the article for important intellectual content: HI, YS, and TA;
(5) Final approval of the article: HI, KS, RE, TA, and MT;
(6) Statistical expertise: HI;
(7) Obtaining funding: HI; and
(8) Data collection and assembly: HI, KS, RE, TA, and MT.

Acknowledgments

This study was supported in part by 1) funding from Omron healthcare Co., Ltd; 2) a Grant-in-Aid from the Japan Society for the Promotion of Science (https://www.jsps.go.jp/) for Research Fellows to HI and 3) JSPS KAKENHI, Grant-in-Aid for Japan Society for the Promotion of Science Young Scientists to HI. The authors thank the members of the Aoyama laboratory (Kyoto University, Kyoto) for assistance with data collection. We also thank Editage (www.editage.jp) for English language editing.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.gaitpost.2018.12.014.

References

[1] A.A. Guccione, D.T. Felson, J.J. Anderson, J.M. Anthony, Y. Zhang, P.W. Wilson, et al., The effects of specific medical conditions on the functional limitations of elders in the Framingham Study, Am. J. Public Health 84 (1994) 351–358.
[2] E.M. Hensor, B. Dube, S.R. Kingsbury, A. Tennant, P.G. Conaghan, Toward a clinical definition of early osteoarthritis: onset of patient-reported knee pain begins on stairs. Data from the osteoarthritis initiative, Arthritis Care Res. 67 (2015) 40–47.
[3] H. Iijima, K. Shimoura, T. Aoyama, M. Takahashi, Biomechanical characteristics of stair ambulation in patients with knee OA: a systematic review with meta-analysis toward a better definition of clinical hallmarks, Gait Posture 62 (2018) 191–201.
[4] T.P. Andriacchi, G.B. Andersson, R.W. Fermier, D. Stern, J.O. Galante, A study of lower-limb mechanics during stair-climbing, J. Bone Joint Surg. Am. 62 (1980) 749–757.
[5] S. Nadeau, B.J. McFadyen, F. Malouin, Frontal and sagittal plane analyses of the stair climbing task in healthy adults aged over 40 years: what are the challenges compared to level walking? Clin. Biomech. (Bristol, Avon) 18 (2003) 950–959.
[6] F. Dobson, R.S. Hinman, E.M. Roos, J.H. Abbott, P. Stratford, A.M. Davis, et al., OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis, Osteoarthr. Cartil. 21 (2013) 1042–1052.
[7] F. Dobson, R.S. Hinman, M. Hall, C.J. Marshall, T. Sayer, C. Anderson, et al., Reliability and measurement error of the Osteoarthritis Research Society International (OARSI) recommended performance-based tests of physical function in people with hip and knee osteoarthritis, Osteoarthr. Cartil. 25 (2017) 1792–1796.
[8] J.J. Tolk, R.P.A. Janssen, C.A.C. Prinsen, D. Latijnhouwers, M.C. van der Steen, S.M.A. Bierma-Zeinstra, et al., The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive, Knee Surg. Sports Traumatol. Arthrosc. (2017).
[9] D.M. Kennedy, P.W. Stratford, J. Wessel, J.D. Gollish, D. Penney, Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty, BMC Musculoskelet. Disord. 6 (2005) 3.
[10] T.E. McAlindon, J.B. Driban, Y. Henrotin, D.J. Hunter, G.L. Jiang, S.T. Skou, et al., OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis, Osteoarthr. Cartil. 23 (2015) 747–760.
[11] R. Soangra, T.E. Lockhart, Agreement in gait speed from smartphone and stopwatch for five meter walk in laboratory and clinical environments, Biomed. Sci. Instrum. 50 (2014) 254–264.
[12] O. Bruyere, C. Cooper, N. Arden, J. Branco, M.L. Brandi, G. Herrero-Beaumont, et al., Can we identify patients with high risk of osteoarthritis progression who will respond to treatment? A focus on epidemiology and phenotype of osteoarthritis, Drugs Aging 32 (2015) 179–187.
[13] K. Shimoura, H. Iijima, Y. Suzuki, T. Aoyama, Immediate effects of transcutaneous electrical nerve stimulation on pain and physical performance in individuals with pre-radiographic knee osteoarthritis: a randomized controlled trial, Arch. Phys. Med. Rehabil. (2018).
[14] M. Akai, T. Doi, K. Fujino, T. Iwaya, H. Kurosawa, T. Nasu, An outcome measure for Japanese people with knee osteoarthritis, J. Rheumatol. 32 (2005) 1524–1532.
[15] J.H. Kellgren, J.S. Lawrence, Radiological assessment of osteo-arthrosis, Ann. Rheum. Dis. 16 (1957) 494–502.
[16] H. Iijima, T. Aoyama, K. Nishitani, H. Ito, N. Fukutani, T. Isho, et al., Coexisting lateral tibiofemoral osteoarthritis is associated with worse knee pain in patients with mild medial osteoarthritis, Osteoarthr. Cartil. 25 (2017) 1274–1281.
[17] H. Iijima, Y. Suzuki, T. Aoyama, M. Takahashi, Interaction between low back pain and knee pain contributes to disability level in individuals with knee osteoarthritis: a cross-sectional study, Osteoarthr. Cartil. (2018).
[18] S. Robbins, E. Waked, Balance and vertical impact in sports: role of shoe sole materials, Arch. Phys. Med. Rehabil. 78 (1997) 463–467.
[19] S.A. Arnadottir, V.S. Mercer, Effects of footwear on measurements of balance and gait in women between the ages of 65 and 93 years, Phys. Ther. 80 (2000) 17–27.
[20] R. Eguchi, A. Yorozu, M. Takahashi, Accessible ground reaction force estimation using insole force sensors without force plates, Control Conference (ASCC), 2017 11th Asian (2017) 2861–2865.
[21] R.C. Gonzalez, A.M. Lopez, J. Rodriguez-Uria, D. Alvarez, J.C. Alvarez, Real-time gait event detection for normal subjects from lower trunk accelerations, Gait Posture 31 (2010) 322–325.
[22] W. Zijlstra, A.L. Hof, Assessment of spatio-temporal gait parameters from trunk accelerations during human walking, Gait Posture 18 (2003) 1–10.
[23] M.H. Pham, M. Elshehabi, L. Haertner, S. Del Din, K. Srulijes, T. Heger, et al., Validation of a step detection algorithm during straight walking and turning in patients with Parkinson's disease and older adults using an inertial measurement unit at the lower back, Front. Neurol. 8 (2017) 457.
[24] A. Kose, A. Cereatti, U. Della Croce, Estimation of traversed distance in level walking using a single inertial measurement unit attached to the waist, Conference Proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2011 (2011) 1125–1128.
[25] A. Mansfield, G.M. Lyons, The use of accelerometry to detect heel contact events for use as a sensor in FES assisted walking, Med. Eng. Phys. 25 (2003) 879–885.
[26] F.A. Storm, C.J. Buckley, C. Mazza, Gait event detection in laboratory and real life settings: accuracy of ankle and waist sensor based methods, Gait Posture 50 (2016) 42–46.
[27] G.J. Almeida, C.A. Schroeder, A.B. Gil, G.K. Fitzgerald, S.R. Piva, Interrater reliability and validity of the stair ascend/descend test in subjects with total knee arthroplasty, Arch. Phys. Med. Rehabil. 91 (2010) 932–938.
[28] H.C. de Vet, C.B. Terwee, D.L. Knol, L.M. Bouter, When to use agreement versus reliability measures, J. Clin. Epidemiol. 59 (2006) 1033–1039.
[29] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, J. Chiropr. Med. 15 (2016) 155–163.
[30] G.H. Goncalves, L.F. Selistre, M. Petrella, S.M. Mattiello, Kinematic alterations of the lower limbs and pelvis during an ascending stairs task are associated with the degree of knee osteoarthritis severity, Knee 24 (2017) 295–304.
[31] K.L. Bennell, M.A. Hunt, T.V. Wrigley, D.J. Hunter, F.J. McManus, P.W. Hodges, et al., Hip strengthening reduces symptoms but not knee load in people with medial knee osteoarthritis and varus malalignment: a randomised controlled trial, Osteoarthr. Cartil. 18 (2010) 621–628.