Japanese Journal of Comprehensive Rehabilitation Science (2025)

Original Article

Minimal Detectable Change in Muscle Strength Measurements Obtained Using a Hand-Held Dynamometer in Patients with Stroke

Shota Itoh, RPT, PhD,[1] Hiroki Tanikawa, RPT, DMSc,[2] Hikaru Kondo, RPT, PhD,[1,3] Sora Ozeki, RPT, MS,[1] Toshiki Ito, RPT,[1] Kenta Fujimura, OTR, DMSc,[2] Toshio Teranishi, RPT, DMSc[2]

[1]Department of Rehabilitation, Fujita Health University Hospital, Toyoake, Japan
[2]Faculty of Rehabilitation, School of Health Sciences, Fujita Health University, Toyoake, Japan
[3]Department of Rehabilitation Medicine, School of Medicine, Fujita Health University, Toyoake, Japan

ABSTRACT
Itoh S, Tanikawa H, Kondo H, Ozeki S, Ito T, Fujimura K, Teranishi T. Minimal Detectable Change in Muscle Strength Measurements Obtained Using a Hand-Held Dynamometer in Patients with Stroke. Jpn J Compr Rehabil Sci 2025; 16: 9−18.

Objective: The current study aimed to evaluate the reliability of muscle strength measurements using a hand-held dynamometer (HHD) in patients with chronic stroke. Further, it examined the minimal detectable change (MDC_{95}).

Methods: Patients who presented with chronic stroke hemiplegia for > 180 days post-stroke onset were analyzed. Muscle strength in the paretic lower limb was assessed using an HHD, and gait speed was evaluated.

Results: For hip flexion, hip adduction, hip abduction, knee extension, ankle dorsiflexion, and ankle plantarflexion, the intra-rater reliability of the muscle strength measurements, as assessed using the intraclass correlation coefficient (ICC), ranged from 0.989 to 0.998. The inter-rater reliability, as assessed using ICC, ranged from 0.886 to 0.939. Bland-Altman analysis did not indicate systematic errors, and the MDC_{95} of each joint movement was calculated. Muscle strength in hip flexion, hip adduction, knee extension, ankle dorsiflexion, and ankle plantarflexion were significantly associated with gait speed, but not with hip abduction strength. The MDC_{95} of each muscle strength measurement was established, thereby providing a criterion for detecting actual changes that exceed the measurement error.

Conclusions: The HHD had a high reliability in measuring lower limb muscle strength in patients with chronic stroke hemiplegia. Moreover, an association was found between individual muscle strength and gait ability. Based on this study, specific target muscles for interventions that aim to improve gait speed can be identified. Further, the use of MDC_{95} allows for a more accurate assessment of the intervention effects.

Key words: Stroke, Hand-held dynamometer (HHD), Muscle strength, Reliability, Minimal detectable change

Correspondence: Shota Itoh, RPT, PhD
Department of Rehabilitation, Fujita Health University Hospital, 1−98 Dengakugakubo, Kutsukake-cho, Toyoake, Aichi 470−1192, Japan.
E-mail: itosho02@fujita-hu.ac.jp
Accepted: March 10, 2025.
Conflict of Interest: The authors declare no conflict of interest.

Introduction

Patients with stroke commonly exhibit gait impairments, including decreased gait independence [1], abnormal gait patterns such as ankle inversion [2], and reduced toe clearance [3]. In patients with stroke, one of the goals of rehabilitation is the restoration of gait ability [4]. In stroke gait rehabilitation, an accurate understanding of the pathology and severity of gait impairments is essential. Various assessments are conducted to evaluate the severity of gait impairments, and gait speed is often used as a representative indicator of gait ability [5]. Decreased gait speed is associated with fall risk, gait independence, and quality of life [6−8]. Hence, gait speed is an important indicator; in fact, it is used as the main outcome measure in several studies.

Gait speed is associated with the severity of motor paralysis [9]. Further, it is more strongly correlated with muscle strength than motor paralysis [10]. Based on a meta-analysis on the effects of strength training in patients with stroke, improvements in paretic lower limb muscle strength were accompanied by increased gait speed [11]. Therefore, to improve gait speed, a detailed assessment of which muscles in the paretic lower limb exhibit weakness should be conducted. Moreover, the appropriate target muscles for intervention must be identified.

Muscle strength assessment is commonly performed using manual muscle testing (MMT). MMT is utilized to assess muscle strength by applying manual resistance, and it is evaluated using a six-grade ordinal scale. It does not require measurement devices, and it allows for a simple and convenient assessment. However, it may not detect actual increases in muscle strength [12], and a grading of 4 and 5 is subjectively examined, thereby leading to a lack of objectivity [13]. The Fugl-Meyer Assessment (FMA), Stroke Impairment Assessment Set (SIAS), and MMT have limitations in providing detailed evaluations of individual muscles. In addition, because these assessments use an ordinal scale, detecting subtle changes remains challenging.

A portable hand-held dynamometer (HHD) is frequently used to quantitatively assess muscle strength. Previous studies have investigated the reliability of HHD measurements in various diseases [14−29]. As an HHD is easy to use, it is widely utilized in clinical settings. According to the Consensus-Based Standards for the Selection of Health Measurement Instruments checklist, studies with a sample size of < 20 are of low quality [30]. According to a systematic review of the reliability of HHD measurements, most studies have a sample size of < 20 [31].

When evaluating treatment effects, a patient's change is considered an actual change only if it exceeds the measurement error. Therefore, identifying the measurement error for each assessment technique and participants is crucial for accurately determining treatment effects. In addition to validating the reliability of muscle strength measurements in patients with stroke, previous reports have not analyzed minimal detectable change (MDC_{95}). Further, in patients with knee osteoarthritis, when evaluating knee extensor strength using an HHD, the measurement error increases as the muscle strength improves [22]. However, there are no studies analyzing the association between the severity of motor paralysis and measurement error (variability in muscle strength measurements) in patients with stroke.

This study aimed to assess paretic lower limb muscle strength using an HHD in > 20 patients with stroke hemiplegia. Further, it investigated the reliability of muscle strength measurements and evaluated the MDC. Finally, it analyzed the effects of motor paralysis severity on measurement error in muscle strength assessment and identified the muscles associated with gait speed.

Methods

1. Participants

The inclusion criteria were patients who presented with stroke hemiplegia for > 180 days post-stroke onset and who were treated as outpatients at Fujita Health University Hospital. The patients who provided informed consent were evaluated from March 2011 to September 2024. The exclusion criteria included patients who had difficulties in performing activities of daily living prior to stroke onset, those with a history of lower limb pain or orthopedic conditions, and those with a higher brain dysfunction or cognitive impairment that prevented them from following instructions. The Research Ethics Committee of Fujita Health University approved this study (HM22-161). Prior to conducting the research, the purpose and details of the experiment were comprehensively explained to the participants both verbally and in writing, and written informed consent was obtained.

2. Equipment and the Research Protocol

Age, sex, disease, affected side, time after onset, and the motor section of the Stroke Impairment Assessment Set (SIAS-m) scores [32] were assessed. In addition, muscle strength in the paretic side was examined for hip flexion, hip adduction, hip abduction, knee extension, ankle plantarflexion, and ankle dorsiflexion. Muscle strength was evaluated using an HHD (μTasF-1, manufactured by Anima Co., Ltd.). Maximum isometric contraction was measured for 3 s, twice consecutively, with an adequate rest interval between each measurement. To verify the intra-rater reliability, one of the two physical therapists with 7 and 6 years of experience conducted the evaluations. To validate the inter-rater reliability, two physical therapists with 9 and 5 years of experience assessed each participant in a randomized order.

The sensor placement was based on the methodology of previous studies [25−29], and measurements were conducted to minimize compensatory movements. Hip flexion, knee extension, and ankle dorsiflexion were examined in the sitting position. Meanwhile, hip adduction, hip abduction, and ankle plantarflexion were evaluated in the supine position (Figure 1). Between the first and second measurements, the sensor was removed from the participants and then repositioned before conducting the next measurement. Before the examinations, the procedure was explained to the participants, and their understanding was confirmed. Sufficient practice trials were conducted before proceeding with the actual measurements. If compensatory movements occurred, the measurement was repeated. Further, to analyze the association between gait ability and muscle strength, the comfortable gait speed was measured while the participants walked barefoot on a 10-m walkway without using assistive devices. In cases where

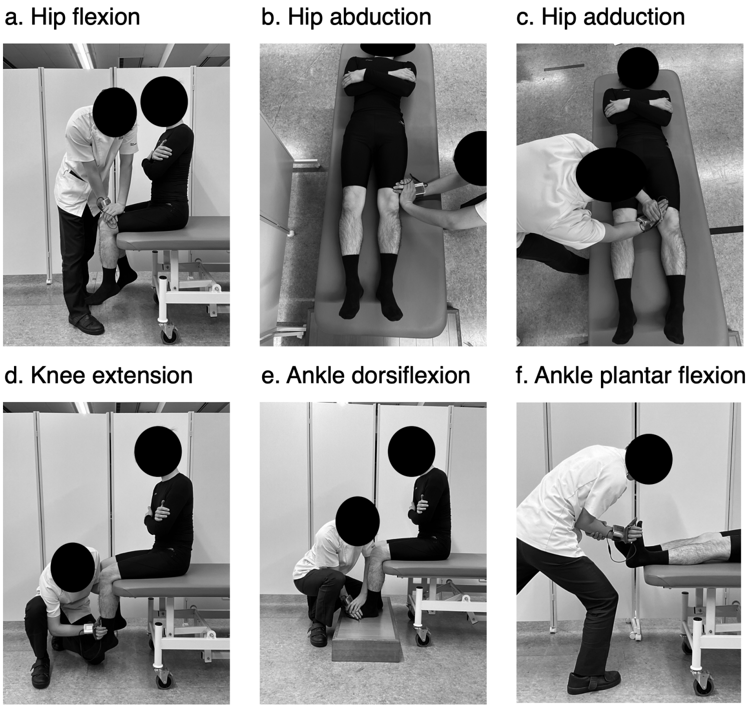

Figure 1. Measurement Posture.
Hip flexion-sensor placement: Distal anterior thigh [25, 26]
Hip abduction-sensor placement: distal lateral thigh [27]
Hip adduction-sensor placement: distal medial thigh [28]
Knee extension-sensor placement: distal anterior lower leg [25, 27]
Ankle dorsiflexion-sensor placement: dorsum of the foot [29]
Ankle plantarflexion-sensor placement: plantar surface of the foot [25, 27]

walking without a cane was difficult, the assessment was conducted with the use of a cane.

3. Analysis

Muscle strength values (N) were calculated by dividing them by each participant's body weight (kg) (N/kg).

3.1 Verification of Intra-rater Reliability

To validate the intra-rater reliability of each joint movement, the ICC (1,2) was calculated using the two measurement values.

3.2 Validation of Inter-rater Reliability

In cases evaluated by two assessors, the average of the two measurements for each assessor was used as the measurement value. To verify the inter-rater reliability of each joint movement, the ICC (2,2) was calculated based on the measurements from both assessors.

3.3 Calculation of MDC

The MDC was calculated using the data obtained from the intra-rater reliability analysis. Spearman's rank correlation coefficient was used to examine the association between SIAS-m scores, which indicate the severity of motor paralysis, and HHD measurements, which represent muscle strength. Hip flexion, hip abduction, and hip adduction were categorized based on SIAS-hip, knee extension based on SIAS-knee, and ankle dorsiflexion and ankle plantarflexion based on SIAS-ankle. Multiple comparisons were performed for each group stratified by SIAS-m scores. The Shapiro-Wilk test was used to assess the normality of the measurement values (N/kg). If normality was not confirmed, the Kruskal-Wallis test was performed. If there was a statistically significant difference, multiple comparisons were conducted using the Steel-Dwass test.

To assess the presence of systematic bias in the obtained measurement values, a Bland-Altman analysis was performed to examine both constant bias and proportional bias. Constant bias was evaluated using Wilcoxon's signed-rank test for the two measurement values. Meanwhile, proportional bias was examined by calculating and analyzing the slope of the regression line. If no systematic bias was detected, the MDC_{95} was calculated. After confirming the presence or absence of systematic bias, the range of error contamination was estimated. If the measurement values obtained from the assessment fell within the estimated error tolerance range, the change was attributed to the measurement error. Conversely, if the values exceeded the estimated error tolerance range, the change was considered a true change.

The MDC is typically evaluated using its 95% confidence interval (MDC_{95}). Thus, it was calculated using the z-value for the 95% confidence interval (1.96) and the standard deviation of the difference (SDd), as shown in the following equation:

$$MDC_{95} = 1.96 \times SD_d \quad \text{(Equation 1)}$$

When calculating the ICC, conducting a Bland-Altman analysis, and calculating the MDC_{95}, patients without joint movement due to severe motor paralysis were excluded from the analysis. When a correlation was observed between the SIAS-m scores and the HHD measurements, the MDC_{95} was also calculated for each SIAS-m score category. To analyze the association between the severity of motor paralysis and measurement error during muscle strength assessment, Spearman's rank correlation coefficient was calculated to examine the association between measurement error and SIAS-hip (hip flexion, hip abduction, and hip adduction), SIAS-knee (knee extension), and SIAS-ankle (ankle dorsiflexion and ankle plantarflexion).

3.4 Association between Muscle Strength and Gait Speed

To analyze the association between muscle strength measurements and gait speed, Spearman's rank correlation coefficient was calculated for the association between gait speed and the average of two measurements for hip flexion, hip adduction, hip abduction, knee extension, ankle dorsiflexion, and ankle plantarflexion.

Statistical analyses were performed using the Statistical Package for the Social Sciences software, version 24 (IBM Inc., Armonk, NY, USA) and JMP 13 (SAS Institute Inc., Cary, NC, USA). A p value of < 0.05 indicated statistically significant differences.

Results

1. Verification of Intra-rater Reliability

The analysis of the intra-rater reliability included 42 cases (Table 1). The intra-rater reliability, which was assessed using the ICC (1,2), was as follows: hip flexion, 0.992; hip abduction, 0.998; hip adduction, 0.990; knee extension, 0.995; ankle dorsiflexion, 0.989; and ankle plantarflexion, 0.991 (Table 3).

2. Validation of Inter-rater Reliability

The analysis of the inter-rater reliability included 12 cases (Table 2). The inter-rater reliability between evaluator 1 and evaluator 2, which was assessed using the ICC (2,2), was as follows: hip flexion, 0.908; hip abduction, 0.874; hip adduction, 0.939; knee extension, 0.932; ankle dorsiflexion, 0.901; and ankle plantarflexion, 0.886 (Table 3).

3. Calculation of MDC

The MDC was calculated using the 42 cases included in the intra-rater reliability analysis. Figure 2 shows the results of the Bland-Altman analysis. No proportional bias or constant bias was observed for any of the items (Table 4). In all participants, the

Table 1. Characteristics of the participants (intrarater reliability).

	Total number of participants (*n* = 42)
Sex (n)	Male: 28, female: 14
Age (years)	56.8 ± 14.5
Height (cm)	163.1 ± 6.9
Weight (kg)	61.0 ± 10.9
Time after onset (days)	1.771 ± 1.972
SIAS-hip (point)	3 (2-3)
SIAS-knee (point)	3 (2-3)
SIAS-ankle (point)	2 (1-3)
Comfortable gait velocity (km/h)	1.29 ± 1.01
Use of a cane (*n*)	With a cane: 3, Without a cane: 35, Unable to walk: 4

SIAS: Stroke Impairment Assessment Set

Table 2. Characteristics of the participants (inter-rater reliability).

	Number of participants (*n* = 12)
Sex (n)	Male: 10, female: 2
Age (year)	54.8 ± 12.9
Height (cm)	167.7 ± 9.4
Weight (kg)	65.6 ± 8.5
Time after onset (days)	2.146 ± 2.206
SIAS-hip (point)	3 (2-4)
SIAS-knee (point)	3 (3-3)
SIAS-ankle (point)	2 (2-2)

SIAS: Stroke Impairment Assessment Set

Table 3. Inter-rater reliability.

	Inter-rater reliability		Intra-rater reliability
	Rater 1 ICC (1, 2)	Rater 2 ICC (1, 2)	ICC (2, 2)
Hip flexion	0.997	0.998	0.908
Hip abduction	0.999	0.993	0.874
Hip adduction	0.999	0.999	0.939
Knee extension	0.999	0.995	0.932
Ankle dorsiflexion	0.998	0.998	0.901
Ankle plantarflexion	0.996	0.998	0.886

ICC: Intraclass correlation coefficient

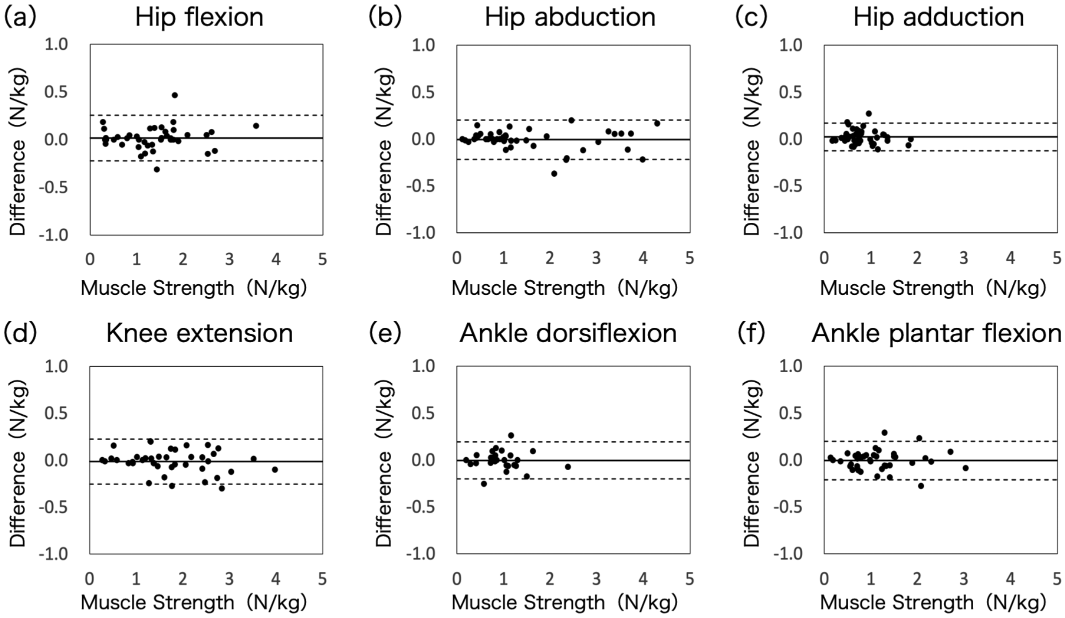

Figure 2. Bland-Altman analysis.
The Bland-Altman plot illustrates the agreement between the first and second measurement values. The black solid line represents the mean difference between the two measurements, and the dotted lines indicate the limits of agreement (LOA).

Table 4. Mean and reliability of the measurements for each joint movement.

	Mean of the measurements (Mean ± standard deviation)			Systematic bias			Reliability	Error
	First (N/kg)	Second (N/kg)	Second-first (N/kg)	Constant bias *p* value	Proportional bias R	Proportional bias *p* value	ICC (1, 2)	MDC_{95}
Hip flexion	1.39 ± 0.76	1.37 ± 0.75	−0.01 ± 0.12	0.83	0.09	0.59	0.992	0.24
Hip abduction	1.55 ± 1.19	1.56 ± 1.21	0.01 ± 0.11	0.80	−0.13	0.41	0.998	0.21
Hip adduction	0.81 ± 0.38	0.79 ± 0.39	−0.02 ± 0.08	0.35	−0.12	0.46	0.990	0.15
Knee extension	1.73 ± 0.87	1.75 ± 0.90	0.02 ± 0.12	0.90	−0.22	0.18	0.995	0.24
Ankle dorsiflexion	0.96 ± 0.47	0.97 ± 0.47	0.01 ± 0.10	0.56	−0.05	0.80	0.989	0.15
Ankle plantarflexion	1.16 ± 0.67	1.17 ± 0.67	0.01 ± 0.10	0.10	0.01	0.93	0.991	0.20

ICC: Intraclass correlation coefficient

MDC_{95} values were 0.24 for hip flexion, 0.21 for hip abduction, 0.15 for hip adduction, 0.24 for knee extension, 0.15 for ankle dorsiflexion, and 0.20 for ankle plantarflexion (Table 4).

4. Association of the Mean Measurement Values and Motor Paralysis and Gait Speed

The mean measurement values were as follows: hip flexion, 1.38 ± 0.75; hip abduction, 1.55 ± 1.10; hip adduction, 0.80 ± 0.38; knee extension, 1.74 ± 0.89;

ankle dorsiflexion, 0.96 ± 0.47; and ankle plantarflexion, 1.17 ± 0.66. For hip flexion, knee extension, ankle dorsiflexion, and ankle plantarflexion, patients with milder motor paralysis had significantly higher muscle strength values, as assessed using SIAS ($p < 0.05$) (Figure 3, Table 5). The correlation coefficients between muscle strength values and walking speed were as follows: hip flexion, 0.68; hip abduction, 0.22; hip adduction, 0.46; knee extension, 0.60; ankle dorsiflexion, 0.62; and ankle plantarflexion, 0.45 (Table 5).

5. Measurement Values and MDC95 based on the Severity of Motor Paralysis

For hip flexion, knee extension, ankle dorsiflexion, and ankle plantarflexion, which were significantly associated with SIAS-m scores, the MDC_{95} was calculated individually for each SIAS-m severity level (Table 6).

6. Association between the Severity of Motor Paralysis and Measurement Error in Each Joint Movement

The correlation efficient between the SIAS-m scores and measurement error during joint movement assessments were as follows: hip flexion, 0.25; hip abduction, –0.24; hip adduction, 0.20; knee extension, 0.18; ankle dorsiflexion, 0.09; and ankle plantarflexion, –0.27 (Table 7).

Discussion

In this study, muscle strength in the paretic lower limb of patients with chronic stroke was examined using an HHD. Moreover, the reliability of muscle

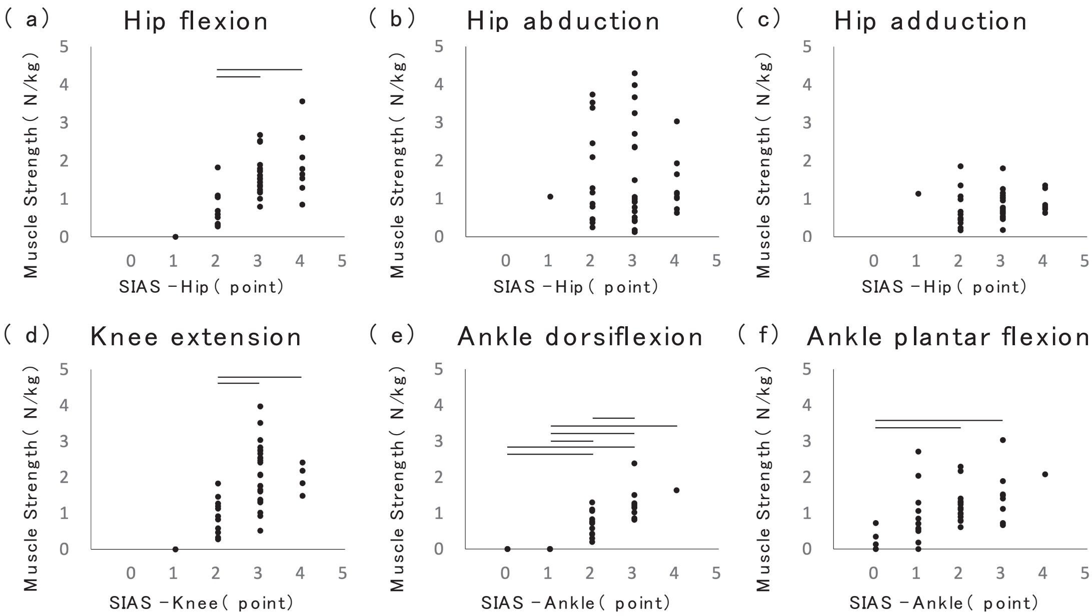

Figure 3. The Steel-Dwass test.
The results of the Steel-Dwass test, which was used to compare muscle strength values across different SIAS-m groups, are presented.

Table 5. Association between the mean measurement values for each joint movement and the severity of motor paralysis and gait velocity.

	Mean (N/kg)	Gait velocity (km/h)	Association with SIAS-m scores		Association with gait velocity	
			Spearman's correlation coefficient	p value	Spearman's correlation coefficient	p value
Hip flexion	1.38 ± 0.75	1.29 ± 1.01	0.65	< 0.01	0.68	< 0.01
Hip abduction	1.55 ± 1.10		0.01	0.98	0.22	0.16
Hip adduction	0.80 ± 0.38		0.24	0.12	0.46	< 0.01
Knee extension	1.74 ± 0.89		0.59	< 0.01	0.60	< 0.01
Ankle dorsiflexion	0.96 ± 0.47		0.90	< 0.01	0.62	< 0.01
Ankle plantarflexion	1.17 ± 0.66		0.46	< 0.01	0.45	< 0.01

The mean was calculated as the average of two consecutive measurements.
The association between the mean of each joint movement and SIAS-m scores as well as the gait speed was analyzed.

Table 6. Mean and MDC_{95} according to the severity of motor paralysis.

<table>
<tr><th></th><th></th><th>SIAS-m 0</th><th>SIAS-m 1</th><th>SIAS-m 2</th><th>SIAS-m 3</th><th>SIAS-m 4</th></tr>
<tr><td>SIAS-hip</td><td>n</td><td>0</td><td>1</td><td>13</td><td>19</td><td>9</td></tr>
<tr><td rowspan="2">Hip flexion</td><td>Mean</td><td>—</td><td>0</td><td>0.67</td><td>1.62</td><td>1.90</td></tr>
<tr><td>MDC₉₅</td><td></td><td>0</td><td>0.30</td><td>0.20</td><td>0.12</td></tr>
<tr><td rowspan="2">Hip abduction</td><td>Mean</td><td>—</td><td>1.05</td><td>1.60</td><td>1.63</td><td>1.36</td></tr>
<tr><td>MDC₉₅</td><td></td><td colspan="4">0.21</td></tr>
<tr><td rowspan="2">Hip adduction</td><td>Mean</td><td>—</td><td>1.13</td><td>0.70</td><td>0.80</td><td>0.87</td></tr>
<tr><td>MDC₉₅</td><td></td><td colspan="4">0.15</td></tr>
<tr><td>SIAS-knee</td><td>n</td><td>0</td><td>2</td><td>11</td><td>25</td><td>4</td></tr>
<tr><td rowspan="2">Knee extension</td><td>Mean</td><td>—</td><td>0</td><td>0.94</td><td>2.06</td><td>1.98</td></tr>
<tr><td>MDC₉₅</td><td></td><td>0</td><td>0.14</td><td>0.28</td><td>0.17</td></tr>
<tr><td>SIAS-ankle</td><td>n</td><td>5</td><td>12</td><td>14</td><td>10</td><td>1</td></tr>
<tr><td rowspan="2">Ankle dorsiflexion</td><td>Mean</td><td>0</td><td>0</td><td>0.73</td><td>1.22</td><td>1.63</td></tr>
<tr><td>MDC₉₅</td><td>0</td><td>0</td><td>0.19</td><td>0.22</td><td>—</td></tr>
<tr><td rowspan="2">Ankle plantarflexion</td><td>Mean</td><td>0.24</td><td>0.93</td><td>1.21</td><td>1.41</td><td>2.08</td></tr>
<tr><td>MDC₉₅</td><td>0.06</td><td>0.23</td><td>0.16</td><td>0.16</td><td>—</td></tr>
</table>

—: indicates that the mean and MDC_{95} could not be calculated due to the absence of relevant cases.
Hip flexion, hip abduction, and hip adduction were analyzed based on the SIAS-hip scores. Knee extension was analyzed based on the SIAS-knee scores, and ankle dorsiflexion and ankle plantarflexion were examined based on the SIAS-ankle scores.
MDC_{95}: minimal detectable change

Table 7. Association between the severity of motor paralysis (SIAS-m) and the variability in the measurement values for each joint movement.

	Spearman's correlation coefficient	*p* value
Hip flexion	0.25	0.11
Hip abduction	–0.24	0.88
Hip adduction	0.20	0.20
Knee extension	0.18	0.28
Ankle dorsiflexion	0.09	0.68
Ankle plantarflexion	–0.27	0.09

The correlation coefficient was calculated to assess the association between the severity of motor paralysis (SIAS-m) and the variability in the measurement values during each joint movement assessment.

strength measurements was evaluated, and the MDC_{95} was calculated. The results showed that the HHD had good reliability for all joint movements, and the MDC for each joint movement was successfully determined.

The criteria for the ICC vary. In some cases, a value of 1 indicates perfect agreement; ≥ 0.9, excellent; ≥ 0.8, good; ≥ 0.7, moderate; ≥ 0.6, acceptable; and < 0.6, reconsideration [33]. In other classifications, an ICC of ≥ 0.8 suggested "almost perfect" [34]. Interpretations may differ. However, an ICC of > 0.75 is required to identify an excellent reliability [35]. In this study, the intra-rater reliability ICC for muscle strength measurements obtained using the HHD in patients with stroke was ≥ 0.99. The inter-rater reliability ICC was ≥ 0.87. Based on these results, the intra-rater reliability and inter-rater reliability were good.

Three factors influence reliability when conducting assessments: subject characteristics, sample size, and the test protocol [36]. The possible causes of variability in muscle strength measurements include instability in the severity of paralysis, poor control of paretic muscles, fatigue, and inadequate understanding of instructions. However, this study targeted patients who presented with chronic stroke for > 180 days. The motor paralysis symptoms of these patients were likely to be stable, thereby reducing variability in muscle output during strength measurements. In addition, muscle fatigue is an important factor of muscle strength measurements. During consecutive strength assessments, incorporating rest periods between evaluations can help decrease the effects of muscle fatigue. In this study, sufficient rest periods were incorporated between measurements and account for muscle fatigue, which likely contributed to improving the reliability of muscle strength assessments. Further, only participants with a good understanding of the instructions were included based on the selection criteria. Thus, measurement errors caused by misunderstanding instructions were considered minimal. Therefore, the measurement error observed in this study was likely less influenced by participant characteristics, and it primarily reflected the measurement error on the evaluator's side.

Studies examining the reliability of assessments suggest that a sample size of at least 30 and, preferably, ≥ 50 is desirable [37,38]. In this study, the sample size for intra-rater reliability verification was 42, which meets the recommended sample size for assessing the reliability. By contrast, the sample size for inter-rater reliability verification was only 12, which was insufficient for a robust reliability assessment. Future studies should include a larger sample size to further validate the findings. To enhance reliability, it is essential to consider and decrease errors that may arise from the test protocol. In this study, to reduce potential errors, all evaluations were conducted by the same evaluator. The participants had a good understanding of the instructions. Thus, errors in task execution were likely minimal during the assessment. Therefore, the measurement error caused by the test protocol was considered minimal. Based on these findings, the likelihood of measurement error was considered low across all factors, including characteristics of the participants, sample size, and test protocol. Therefore, the variability observed between the two trials in this study is likely to reflect fluctuations in the measured parameters rather than systematic error.

To examine whether the measurement error varies with the severity of motor paralysis, the association between the motor paralysis severity and the measurement error was assessed. As a result, no such association was found for any of the assessed muscles. These findings indicate that MDC_{95}, as calculated in this study, may be a useful indicator for detecting changes in muscle strength, regardless of the severity of motor paralysis.

The evaluation of functional impairment in stroke can be categorized into two types: assessments based on movement patterns and assessments focusing on muscle weakness [39]. The evaluation of movement patterns was conducted using ordinal scales such as FMA and SIAS-m. However, clinically, patients with similar FMA or SIAS-m scores commonly exhibit different levels of muscle strength. By contrast, muscle weakness can be assessed using MMT or HHD.

In this study, muscle strength was quantitatively assessed using an HHD. Moreover, it was confirmed that patients with milder motor paralysis for most joint movements had greater muscle strength. Based on these findings, HHD measurements reflect the severity of motor paralysis. However, for hip abduction and hip adduction movements, no association was found between SIAS-m scores, which can reflect hip motor paralysis, and muscle strength. Therefore, individual muscle strength should be examined individually using an HHD, in addition to SIAS, to obtain a more detailed evaluation.

The muscle strength measurements varied within the same SIAS-m score. In addition, the MDC_{95} values calculated in this study were lower than the differences in the mean values across SIAS-m. These findings indicate that even when no changes are detected using ordinal scales such as SIAS-m, detailed muscle strength assessments using an HHD may enable the detection of subtle changes in muscle strength. Further, because the MDC_{95} varied based on the severity of motor paralysis, the degree of paralysis when referencing MDC_{95} should be considered. The MDC_{95} values calculated in this study may be a useful reference for evaluating the effects of therapeutic interventions on the paretic lower limb in patients with chronic stroke.

Further, to analyze which muscle strength is associated with gait ability, the correlation between barefoot comfortable gait speed and muscle strength was examined. The results showed that hip flexion, hip adduction, knee extension, ankle dorsiflexion, and ankle plantarflexion were significantly associated with gait speed. However, there was no association between hip abduction strength and gait speed. According to a review on the association between lower limb muscle strength and gait speed in patients with stroke [40], gait speed is associated with the strength of all lower limb muscles. Further, the study showed that improving overall lower limb muscle strength is essential for enhancing gait speed. However, the degree of association between muscle strength and gait speed varies according to severity. In this study, a significant association between gait speed and muscle strength was observed for all muscle groups except hip abduction, thereby supporting the findings of previous research. The results indicate that muscle strength in hemiplegic patients with motor paralysis is associated with the severity of paralysis. Moreover, a greater muscle strength corresponds to a higher gait ability (speed). By conducting a detailed assessment of which muscles exhibit weakness, it is possible to identify target muscles for interventions that aim to improve gait ability. In addition, the MDC_{95} was determined for each muscle. Thus, it may be a useful reference for evaluating the efficacy of treatments for the muscles affected by weakness.

This study had several limitations. First, all muscle strength assessments were conducted on the same day. Thus, although sufficient rest periods were provided, the potential influence of muscle fatigue could not be ruled out. However, the assessments were conducted on the same day to ensure that measurements were taken during a period when no substantial changes in muscle strength were expected. Second, because this study targeted patients with chronic stroke, the findings may not be directly applicable to muscle strength assessments in patients with acute or subacute stroke. Third, the measurement conditions of gait speed were not completely standardized. There was variability in the use of an assistive device (e.g., a cane) and the inclusion of participants with gait difficulties, leading to a lack of uniformity in gait ability among participants. This might have resulted in inconsistencies in the data on gait speed. To obtain more reliable data, the gait ability conditions during the assessments should be standardized. Future studies should be conducted to evaluate the gait speed under more uniform conditions across participants. Fourth, this study did not include measurements of the hip extensor and knee flexor muscles. Due to time constraints and feasibility considerations, only muscles that could be assessed in the seated and supine positions were measured. Incorporating hip extensor and knee flexor strength assessments in future research could enable a more comprehensive and significant evaluation of lower limb muscle function. Fifth, the sample size for inter-rater reliability analysis was limited. It is generally recommended that at least 30 samples be used to ensure a robustly reliable assessment. However, in this study, the inter-rater reliability was examined with only 12 cases, which may not be sufficient for a conclusive evaluation. Future studies should aim to increase the sample size to further validate inter-rater reliability. In addition, we plan to expand the study to include patients with acute and subacute stroke for further investigation.

Conclusions

In this study, muscle strength in the paretic lower limb of patients with chronic stroke was measured using an HHD. Further, the reliability of muscle strength measurements was examined, and the MDC_{95} was calculated. Finally, the associations between motor paralysis severity and measurement error in muscle strength assessment and between gait speed and muscle strength were analyzed. The results showed no association between motor paralysis severity and measurement error based on the muscle strength assessments. Moreover, the HHD measurements had a good reliability for all measured muscles, thereby allowing the calculation of MDC_{95}. By assessing individual muscle strength and identifying factors contributing to reduced gait ability, this study can help in selecting appropriate treatment strategies. Further, the MDC_{95} can be a useful reference for evaluating the efficacy of interventions targeting specific muscles.

Ethics declaration

This study was conducted in accordance with the ethical principles of the Helsinki Declaration, and after obtaining informed consent from each subject. The study was approved by the Research Ethics Committee of Fujita Health University (HM22-161).

Acknowledgment

We would like to thank the members of Fujita Health University Hospital who helped with the patient assessment.

References

1. Jørgensen HS, Nakayama H, Raaschou HO, Olsen TS. Recovery of walking function in stroke patients: the Copenhagen Stroke Study. Arch Phys Med Rehabil 1995; 76(1): 27–32.
2. Tanikawa H, Kagaya H, Saitoh E, Ozaki K, Hirano S, Itoh N, et al. Efficacy of Botulinum Toxin A Treatment for Pes Varus during Gait. J Stroke Cerebrovasc Dis 2015; 24(10): 2416–22.
3. Matsuda F, Mukaino M, Ohtsuka K, Tanikawa H, Tsuchiyama K, Teranishi T, et al. Biomechanical factors behind toe clearance during the swing phase in hemiparetic patients. Top Stroke Rehabil 2017; 24(3): 177–82.
4. Jones PS, Pomeroy VM, Wang J, Schlaug G, Tulasi Marrapu S, Geva S, et al; SWIFT-Cast investigators. Does stroke location predict walk speed response to gait rehabilitation? Hum Brain Mapp 2016; 37(2): 689–703.
5. Salbach NM, Guilcher SJ, Jaglal SB. Physical therapists' perceptions and use of standardized assessments of walking ability post-stroke. J Rehabil Med 2011; 43(6): 543–9.
6. Perry J, Garrett M, Gronley JK, Mulroy SJ. Classification of walking handicap in the stroke population. Stroke 1995; 26(6): 982–9.
7. Schmid A, Duncan PW, Studenski S, Lai SM, Richards L, Perera S, et al. Improvements in speed-based gait classifications are meaningful. Stroke 2007; 38(7): 2096–100.
8. Khanittanuphong P, Tipchatyotin S. Correlation of the gait speed with the quality of life and the quality of life classified according to speed-based community ambulation in Thai stroke survivors. NeuroRehabilitation 2017; 41(1): 135–41.
9. Bowden MG, Balasubramanian CK, Neptune RR, Kautz SA. Anterior-posterior ground reaction forces as a measure of paretic leg contribution in hemiparetic walking. Stroke 2006; 37(3): 872–6.
10. Hsu AL, Tang PF, Jan MH. Analysis of impairments influencing gait velocity and asymmetry of hemiplegic patients after mild to moderate stroke. Arch Phys Med Rehabil 2003; 84(8): 1185–93.
11. Wist S, Clivaz J, Sattelmayer M. Muscle strengthening for hemiparesis after stroke: A meta-analysis. Ann Phys Rehabil Med 2016; 59(2): 114–24.
12. Schwartz S, Cohen ME, Herbison GJ, Shah A. Relationship between two measures of upper extremity strength: manual muscle test compared to hand-held myometry. Arch Phys Med Rehabil 1992; 73(11): 1063–8.
13. Dvir Z. Grade 4 in manual muscle testing: the problem with submaximal strength assessment. Clin Rehabil 1997; 11(1): 36–41.
14. Bohannon RW. Hand-held compared with isokinetic dynamometry for measurement of static knee extension torque (parallel reliability of dynamometers). Clin Phys Physiol Meas 1990; 11(3): 217–22.
15. Bohannon RW. Test-retest reliability of hand-held dynamometry during a single session of strength assessment. Phys Ther 1986; 66(2): 206–9.
16. Riddle DL, Finucane SD, Rothstein JM, Walker ML. Intrasession and intersession reliability of hand-held dynamometer measurements taken on brain-damaged patients. Phys Ther 1989; 69(3): 182–94.
17. Bohannon RW, Andrews AW. Interrater reliability of hand-held dynamometry. Phys Ther. 1987; 67(6): 931–3.
18. Tanikawa H, Mukaino M, Matsuda F, Inagaki K, Ohtsuka K, Kagaya H, et al. Influence of contralateral lower limb stabilization on hip abductor muscle strength measured by Hand-Held Dynamometer. Jpn J Compr Rehabil Sci 2015; 6: 137–42.

19. Mentiplay BF, Perraton LG, Bower KJ, Adair B, Pua YH, Williams GP, et al. Assessment of Lower Limb Muscle Strength and Power Using Hand-Held and Fixed Dynamometry: A Reliability and Validity Study. PLoS One 2015; 10(10).
20. Wong SS, Yam MS, Ng SS. The Figure-of-Eight Walk test: reliability and associations with stroke-specific impairments. Disabil Rehabil 2013; 35(22): 1896–902.
21. Aguiar LT, Martins JC, Quintino LF, de Brito SAF, Teixeira-Salmela LF, de Morais Faria CDC. A Single Trial May Be Used for Measuring Muscle Strength With Dynamometers in Individuals With Stroke: A Cross-Sectional Study. PM R 2019; 11(4): 372–8.
22. Chopp-Hurley JN, Wiebenga EG, Gatti AA, Maly MR. Investigating the Test-Retest Reliability and Validity of Hand-Held Dynamometry for Measuring Knee Strength in Older Women with Knee Osteoarthritis. Physiother Can 2019; 71(3): 231–8.
23. Thorborg K, Bandholm T, Hölmich P: Hip- and knee-strength assessments using a hand-held dynamometer with external belt-fixation are inter-tester reliable. Knee Surg Sports Trau- matol Arthrosc 2013; 21: 550–5.
24. Buckinx F, Croisier JL, Reginster JY, Dardenne N, Beaudart C, Slomian J, et al. Reliability of muscle strength measures obtained with a hand-held dynamometer in an elderly population. Clin Physiol Funct Imaging 2017; 37(3): 332–40.
25. Mentiplay BF, Tan D, Williams G, Adair B, Pua YH, Bower KJ, et al. Assessment of isometric muscle strength and rate of torque development with hand-held dynamometry: Test-retest reliability and relationship with gait velocity after stroke. J Biomech 2018; 75: 171–5.
26. Arnold CM, Warkentin KD, Chilibeck PD, Magnus CR. The reliability and validity of handheld dynamometry for the measurement of lower-extremity muscle strength in older adults. J Strength Cond Res 2010; 24(3): 815–24.
27. Aguiar LT, Camargo LBA, Estarlino LD, Teixeira-Salmela LF, Faria CDCM. Strength of the lower limb and trunk muscles is associated with gait speed in individuals with sub-acute stroke: a cross-sectional study. Braz J Phys Ther 2018; 22(6): 459–66.
28. Daubney ME, Culham EG. Lower-extremity muscle force and balance performance in adults aged 65 years and older. Phys Ther 1999; 79(12): 1177–85.
29. Katoh M, Yamasaki H: Test-retest reliability of isometric leg muscle strength measurements made using a hand-held dynamometer restrained by a belt: comparisons during and between sessions. J Phys Ther Sci 2009; 21(3): 239–44.
30. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012; 21(4): 651–7.
31. Chamorro C, Armijo-Olivo S, De la Fuente C, Fuentes J, Javier Chirosa L. Absolute Reliability and Concurrent Validity of Hand Held Dynamometry and Isokinetic Dynamometry in the Hip, Knee and Ankle Joint: Systematic Review and Meta-analysis. Open Med (Wars) 2017; 12: 359–75.
32. Chino N, Sonoda S, Domen K, Saitoh E, Kimura A. Stroke impairment assessment set (SIAS): a new evaluation instrument for stroke patients. Jpn J Rehabil Med 1994; 31: 119–25.
33. Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. Clin Exp Pharmacol Physiol 2002; 29(7): 527–36.
34. Shrout PE: Intraclass Correlations: uses in assessing rater reliability. Psychol Bull 1979; 86(2): 420–8.
35. Fleiss JL: The Design and Analysis of Clinical Experiments. Wiley, New York - Chichester - Brislane - Toronto - Singapore 1986, 432 S.
36. Flansbjer UB, Holmbäck AM, Downham D, et al.: Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med 2005; 37(2): 75–82.
37. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000; 30(1): 1–15.
38. Donner A, Eliasziw M. Sample size requirements for reliability studies. Stat Med 1987; 6(4): 441–8.
39. Saitoh E, Chino N. Physical function assessments for patients with stroke impairments. Sogo Rehabil 1989; 11: 481–9.
40. Dorsch S, Ada L, Sorial T, Fanayan E. The Relationship Between Strength of the Affected Leg and Walking Speed After Stroke Varies According to the Level of Walking Disability: A Systematic Review. Phys Ther 2021; 101(12).