What change in isokinetic knee muscle strength can be detected in men and women with hemiparesis after stroke?

Flansbjer, Ulla-Britt; Drake, Anna Maria; Downham, David; Lexell, Jan

Published in:
Clinical Rehabilitation

DOI:
10.1191/0269215505cr854oa

2005

Flansbjer, Ulla-Britt and Holmbäck, Anna Maria and
Downham, David and Lexell, Jan

"What change in isokinetic knee muscle strength can be detected in men
and women with hemiparesis after stroke?"

Clin Rehabil. 2005 Aug;19(5):514-22.

Publisher: Arnold Publishers

What change in isokinetic knee muscle strength can be detected in men and women with hemiparesis after stroke?

Ulla-Britt Flansbjer Department of Rehabilitation, Lund University Hospital, Lund, Department of Health Sciences, Lund University, Malmö, Sweden, **Anna Maria Holmbäck** Department of Physical Therapy, Lund University Hospital, Lund, Sweden, **David Downham** Department of Mathematical Sciences, University of Liverpool, Liverpool, England and **Jan Lexell** Department of Rehabilitation, Lund University Hospital, Lund, Department of Health Sciences, Lund University, Malmö, Department of Health Sciences, Luleå, University of Technology, Boden, Sweden

Address for correspondence: Ulla-Britt Flansbjer, Department of Rehabilitation, Lund University Hospital, Orupssjukhuset, 221 85 Lund, Sweden.

Phone: (+46) 413 55 67 48. Fax: (+46) 413 55 67 09.

E-mail: ulla-britt.flansbjer@skane.se

Running title: Reliability of muscle strength measurements in stroke

ABSTRACT

Objective: To assess the intra-rater (between occasions) test-retest reliability of isokinetic knee muscle strength measurements in subjects with chronic post-stroke hemiparesis and to define limits for the smallest change that indicates real (clinical) improvements for stroke patients.

Subjects: Fifty men and women (mean age 58 ±6.4 years) 6 to 46 months post-stroke, able to walk at least 300 m with or without a unilateral assistive device.

Methods: Maximal concentric knee extension and flexion contractions at 60º/s and 120º/s, and maximal eccentric knee extension contractions at 60º/s, with the paretic and non-paretic limbs, were performed seven days apart using a Biodex dynamometer.

Measures: Reliability of the maximum peak torque measurements was evaluated with the intraclass correlation coefficient ($ICC_{2,1}$), the Bland & Altman analyses, the standard error of measurement (SEM and SEM%) and the smallest real difference (SRD and SRD%).

Results: Test-retest agreements were high ($ICC_{2,1}$ 0.89-0.96) with no discernible systematic differences between limbs, angular velocities and modes. The standard error of measurement (SEM%), representing the smallest change that indicates a real (clinical) improvement for a group of subjects, was relatively small (8-20%). The smallest real difference (SRD%), representing the smallest change that indicates a real improvement for a single subject ranged from 26% to 33% for concentric knee extension, 39% to 55% for concentric knee flexion, and 22% to 25% for eccentric knee extension.

Conclusion: Isokinetic knee muscle strength can be reliably measured and used to detect real improvements following an intervention for single subjects as well as for groups of subjects with chronic mild to moderate hemiparesis after stroke.

Introduction

Hemiparesis with reduced motor function is one of the main impairments after stroke. Current evidence suggests that reduced muscle strength may be responsible for the compromised motor function.[1] Several studies have shown that muscle strength in subjects after stroke is significantly correlated with, for example, gait speed and postural sway.[1] These findings have lead to an increased interest in resistance training of the lower limbs as a way of improving muscle strength in subjects with post-stroke hemiparesis, particularly after the acute phase.[2]

To evaluate changes following strength training in stroke, we need equipment and methods that provide reliable measurements of muscle strength. Reliability is a broad concept that incorporates the agreement between measurements, the presence of systematic changes in mean and the size of measurement errors.[3] Several statistical methods and indices are required to fully assess the reliability of measurements.[4-10] If the equipment and methods can be shown to be sufficiently sensitive to detect real changes in muscle strength, then real (clinical) improvements could be indicated whenever measurements following an intervention lie outside the range of 'measurement noise'. Results from the reliability analysis can be used to define limits for the smallest changes in muscle strength that indicate such improvements, both for a group of stroke patients and for individual stroke patients.

Isokinetic dynamometers are frequently used to assess muscle strength in both health and disease. They enable the measurement of muscular moment under concentric as well as eccentric actions and at different angular velocities, and so provide information about muscle performance related to functional tasks.[11] Only a few studies have evaluated the reliability of isokinetic concentric muscle strength measurements in chronic stroke patients (more than 6 months following stroke-onset).[12-15] Basing their arguments on the intraclass correlation coefficient (ICC), the measurements were generally considered to be reliable.

None of these studies addressed systematic changes in the mean and measurement errors, nor did they define limits for the smallest changes that indicate real improvements. Moreover, no study has specifically evaluated the reliability of eccentric muscle strength measurements in stroke patients. Thus, further studies of the reliability of isokinetic muscle strength measurements in stroke patients are needed.

The overall aims of this study were i) to assess the reliability of isokinetic concentric and eccentric muscle strength measurements in subjects with mild to moderate chronic post-stroke hemiparesis and ii) to define limits for the smallest change that indicate real (clinical) improvements in muscle strength following an intervention, both for a group of stroke patients and for individual stroke patients.

Methods

Subjects

A sample of 50 community-dwelling subjects (38 men and 12 women) was recruited from the Comprehensive Integrated Rehab Unit database in the Department of Rehabilitation, Lund University Hospital. As a general principle, the larger the sample size the more compelling is the argument for extrapolating the measurement method to a given population. It has therefore been suggested that sample sizes of at least 30 subjects should be considered in reliability studies.[9]

Prior to the first test session, all subjects completed a questionnaire that provided demographic and medical information, and were medically checked. The ages for the men were [mean ±SD, range] 59 ±7, 46-72 years, and for the women 58 ±5, 50-66 years. The times from the ischemic or hemorrhagic stroke onset until the first test session were 16 ±5, 6-46 months, for the men and 18 ±5, 6-33 months, for the women. Every subject met the following inclusion criteria: i) a minimum of 6 months and a maximum of 48 months post-stroke; ii) discharged from interdisciplinary rehabilitation services; iii) residual hemiparesis at discharge from primary rehabilitation; iv) at the time of the study, the ability to walk at least 300 m with or without a unilateral assistive device and to understand both verbal and written information; and v) medically stable with no other diseases that significantly influenced muscle strength. All subjects were also participating in a study of the reliability of gait performance tests.[6]

Ethics

All subjects were contacted by phone, received written information and thereafter gave their informed consent. At the first test-session each subject was again informed about the purpose and disposition of the study. The Ethics Research Committee of Lund University, Lund, Sweden approved the study.

Pre-test assessments

To characterize the group, each subject was interviewed and scored with the Functional Independence Measure motor domain (FIM; Swedish version of FIM^{SM})[16] prior to the first test session. A majority (94%) of the subjects were rated independent/modified independent by the FIM, i.e., a FIM motor domain score equal to or greater than 78. The occurrence of spasticity in the paretic leg was assessed with the Modified Ashworth Scale (MAS)[17] before each of the two test sessions. The MAS is a 6-point rating scale, ranging from 0 (no increase in tone, both low and normal tone) to 5 (the limb is rigid in flexion or extension). The subjects' hip, knee and ankle joints were tested in a supine position with shoes and ankle-foot orthosis removed. Most of the subjects had very little or no spasticity; 19 subjects were scored 0 on both occasions on the MAS and only 8 subjects were scored more than 3 points on each occasion.

Equipment and positioning

Measurements were performed on a Biodex® Multi-Joint System II isokinetic dynamometer (Biodex Medical Systems, Inc., Shirley, New York, USA) with the Biodex Advantage software version 4.0. The standard Biodex knee unit attachment was used. Before each test session, the system was calibrated to be within allowable limits set by the Biodex.

Each subject was seated in the adjustable chair of the dynamometer, without any orthosis. The test position was sitting with back support and hip flexion 85º. The dynamometer was positioned so that the lateral knee joint line was aligned with the movement axis of the dynamometer. The subject was firmly stabilised with straps across the shoulders, waist and thigh. The ankle cuff of the lever arm was strapped 3 cm proximal to the malleoli of the tested limb. The subjects sat with folded arms throughout the test to minimize upper extremity involvement. Before each measurement the full range of motion (ROM) was set. To account for the influence of the gravity effect torque on the data, each subject's limb was weighed and the Biodex software corrected the data. The details of the position were recorded and used in the second test session.

Knee muscle strength measurements

Each subject was tested on two occasions at the same time of the day. For 49 subjects the interval between the two test sessions was 7 days and for one subject 14 days. All subjects were provided transport free of charge to and from the test site. The same physiotherapist (U-BF) performed all measurements. Each test session lasted approximately one hour and the subjects were offered refreshments (water or apple juice) during the sessions.

To warm up and become familiar with the Biodex equipment and the test procedure, each subject performed 10 reciprocal knee extension and flexion contractions of the non-paretic limb with no resistance, followed by three contractions with some resistance and ending with two maximal contractions. After a 1-minute rest, each subject performed in succession three maximal concentric extension and flexion contractions of the knee at 60º/s with the non-paretic limb. Following a 2-minutes rest, testing continued with five concentric extension and flexion contractions of the knee at 120º/s. After a further 5-minutes rest the same warm up, familiarization and test procedure was performed with the paretic limb. Throughout the tests, subjects were asked to push and pull as hard and fast as possible. The concentric measurements ended with a 10-minutes rest.

For the eccentric strength measurements, the ROM was set from 95º for knee flexion to between 20º and 30º of flexion for knee extension. By doing an active flexion, subjects initiated the shaft movement of the Biodex dynamometer and the limb was lifted to the extended position. The eccentric contraction measurements started when subjects made a forceful extension in the dynamometer. The torque limits were adjusted and set for the test. The subjects tested the technique on the non-paretic side by performing 5 to 10 submaximal and 5 maximal contractions. After a 2-minutes rest, the subjects performed in succession, five maximal eccentric knee extension contractions at 60°/s with the non-paretic limb. After a 5-minutes rest the same familiarization and test procedure was performed with the paretic limb. The instructions for the eccentric strength measurements were: “make a slight knee flexion and let the machine lift your leg, extend your leg against the cuff and resist forcefully the pressure from the lever arm throughout the whole range of movement”.

The subjects were not allowed to see the torque display during the measurements, but consistent verbal encouragement was given throughout. A written summary and oral information about the test results were given at the end of the second test session.

All subjects were able to perform the concentric contractions at 60º/s, with both the paretic and non-paretic limbs, and at 120º/s with the non-paretic limb. Two men were unable to perform the concentric extension and flexion contractions at 120º/s with the paretic limb and six subjects (another man and three women) were unable to perform the eccentric knee extension contractions at 60º/s with the paretic limb. The peak torque values revealed that these six subjects performed significantly less well than the remaining 44 subjects at 60º/s and 120º/s for knee flexion with the paretic limb (Mann-Whitney test: p=0.029 and p=0.033, respectively), whereas the MAS and the FIM were not significantly different between these two groups. The statistical analyses are based on all 50 subjects except for 120º/s knee extension and flexion of the paretic limb (n=48), and for eccentric knee extension (n=44).

Data and statistical analyses

The Biodex evaluation report provided data from each angular velocity and mode from the paretic and non-paretic limbs. The contractions with the highest peak torque value (Newton meter; Nm) from each of the two test sessions were used in the reliability analyses. All calculations were performed using the SPSS 11.0 Software for Windows (SPSS Inc., Chicago, Ill., USA). The men were generally stronger than the women, but as there were no other discernible systematic differences between the sexes, data for the men and the women were combined throughout the analyses and presentations. The statistical methods described here were also used in our study of the reliability of gait performance tests in chronic stroke patients.[6]

Agreement between measurements was assessed by the intraclass correlation coefficient, $ICC_{2,1}$.[18] If BMS represents the variability between subjects, WMS the variability in the measurements within subjects, JMS the variability between test sessions, EMS the residual mean square, and n the number of subjects, then for two test sessions

$$ICC_{2,1} = (BMS-EMS) / (BMS+EMS+2(JMS-EMS)/n) \qquad [1]$$

A two-way ANOVA (random effect model) was used and the 95% confidence interval (95% CI) for $ICC_{2,1}$ was obtained from the ANOVA tables. The $ICC_{2,1}$ was used in this study as this form is applicable when all subjects are tested twice by the same rater.[18]

Systematic changes in the mean were assessed with the 'Bland & Altman analyses'.[5] The 'Bland & Altman analyses' included the calculation of

$\bar{d}$ = the mean difference between the two test sessions [2]

SD_{diff} = the standard deviation of the differences between the two test sessions [3]

standard error (SE) of $\bar{d} = SD_{diff} / \sqrt{n}$ [4]

95% confidence intervals of $\bar{d}$ (95% CI) = $\bar{d} \pm 2.01 \times SE$ [5]

The value 2.01 in equation [5] was obtained from the t-table with 49 and 47 (n-1) degrees of freedom (df); for 43 degrees of freedom the value is 2.02. If zero is included within the 95% CI, it is inferred that there is no significant systematic bias in the data. The 'Bland & Altman analyses' also included the formation of the 'Bland & Altman graphs', in which the difference between test sessions 2 and 1 (2 minus 1) is plotted against the mean of the two test sessions for each subject. These graphs illustrate systematic variations around the zero line and heteroscedasticity, which occurs when the difference between test-retest measurements increase as the mean value of the measurements increase. The possibility of heteroscedasticity was addressed by forming the Pearson's correlation coefficient of the absolute differences between test sessions 2 and 1 and the means of the two test sessions.[19]

Measurement errors were evaluated by the standard error of measurement, SEM, and the SEM%. The SEM was calculated using the square root of the within-subjects error variance

$$SEM = \sqrt{WMS} \qquad [6]$$

The SEM%, the within subject standard deviation as a percentage of the mean, was defined by

$$SEM\% = (SEM/mean) \times 100 \qquad [7]$$

where mean is the mean for all the observations from test sessions 1 and 2. The SEM% is independent of the units of measurement, and represents the limit for the smallest change that indicates a real improvement for a group of subjects following, for example, an intervention.

To define the smallest change that indicates a real improvement or a deterioration for a single subject, we used the smallest real difference, SRD, introduced by Beckerman et al.[20] The SRD was defined by

$$SRD = 1.96 \times SEM \times \sqrt{2} \qquad [8]$$

An 'error band' around the mean difference of the two measurements, $\bar{d}$, was defined by

$$95\%\ SRD = \bar{d} \pm SRD \qquad [9]$$

The SRD can be expressed as a percentage value, SRD%, which is independent of the units of measurement (in analogy with SEM%). The SRD% was defined by

$$\text{SRD\%} = (\text{SRD}/\text{mean}) \times 100 \qquad [10]$$

where mean is the mean for all observations from test sessions 1 and 2.[6]

Results

Using the criteria of Fleiss,[21] all peak torque measurements showed excellent agreement; the values of $ICC_{2,1}$ ranged from 0.89 to 0.96 (Table 1). The 95% confidence intervals for $ICC_{2,1}$ were narrow ranging from 0.81 to 0.98.

Insert Table 1 about here

All $\bar{d}$ values were low and the widths of the 95% CI for $\bar{d}$ were narrow (Table 1). The positive value of $\bar{d}$ for all the concentric peak torque measurements indicated that the performance at the second test session was generally better than at the first. For five of the eight concentric measurements, zero was not included in the 95% CI of $\bar{d}$ implying a significantly ($p<0.05$) different performance between the two test sessions. For the eccentric peak torque measurements, the $\bar{d}$ value indicated that the performance was generally better at the first test session; here, zero was included in the 95% CI of $\bar{d}$

implying that the difference between the two test sessions for the eccentric peak torque measurements was not significant ($p>0.05$).

From the ‘Bland & Altman graphs’ (Figure 1), there were generally more concentric values above the zero line than below, illustrating the significantly better performance at the second test session for five of the eight concentric measurements (see previous paragraph). There was also significantly ($p<0.05$) larger variability for higher test values, i.e. heteroscedasticity, from the paretic limb in three of the four concentric measurements. The Pearson’s correlation coefficient of the absolute differences between test sessions 2 and 1 and the mean of the two test sessions for each subject for these measurements were $r=0.31$ (60°/s knee extension), $r=0.41$ (120°/s knee extension) and $r=0.54$ (120°/s knee flexion).

Insert Figure 1 about here

The SEM gives the measurement errors in absolute values and represents the limit for the smallest change that indicates a real (clinical) improvement for a group of subjects following, for example, an intervention. The SEM% is independent of the units of measurement; the values of SEM% ranged from 7.8% to 20.0% (Table 1).

The 95% SRD represents the limits for the smallest change that indicates a real (clinical) change for a single subject. The SRD%, which represents the change in relative terms, ranged from 26% to 33% for concentric knee extension, from 39% to 55% for concentric knee flexion, and from 22% to 25% for eccentric knee extension (Table 1).

Discussion

The main finding in this study was that isokinetic knee muscle strength could be reliably assessed in both the paretic and non-paretic limbs. This supports the use of isokinetic dynamometers for the assessment of knee muscle strength in subjects with chronic mild to moderate post-stroke hemiparesis and makes it possible to detect changes that indicate real (clinical) improvements.

The ICC values for all the concentric measurements were high and comparable with ICC values from healthy subjects.[22, 23] Previous studies of the test-retest reliability of concentric knee extension peak torque measurements at 30º/s to 120º/s in chronic stroke patients (n=9 to 20) have also found high ICC values (0.81 to 0.99).[12-15] In three of these studies the reliability of concentric knee flexion peak torque measurements (60º/s and 120º/s) were evaluated. Two studies[12, 14, 15] reported high ICC values (0.91 to 0.96; 60º/s and 120º/s), whereas one[14] reported a high ICC value (0.90) for the less-affected limb but a low value (0.48) for the affected limb (60º/s). The ICC values for knee flexion in the present study were slightly lower than for knee extension, but only for the non-paretic limb. Plausible explanations for the different results are differences in the degree of post-stroke muscle weakness and in the sample size.

No study, to the best of our knowledge, has specifically evaluated the test-retest reliability of isokinetic eccentric strength measurements in chronic stroke patients. The ICC values for these measurements were found to be as high as for the concentric measurements, and comparable with ICC values from healthy subjects.[22, 24]

There were indications of a learning effect for all concentric measurements but not for the eccentric measurements. The $\bar{d}$ values were positive for all concentric measurements and for five of the eight measurements, zero was not included in the 95% CI of $\bar{d}$ indicating a significantly better performance during the second test session (cf. Table 1). The mean differences between the two sessions ($\bar{d}$) were, however, small compared with the mean for the two test sessions (cf. Figure 1) and the 95% CI of $\bar{d}$ were narrow (cf. Table 1), implying that this learning effect was small. Eng et al.[12] also reported a learning effect for knee extension and flexion, and that the differences between the two test sessions were small and not significant. Possible learning effects might be reduced if practice sessions are allowed before the actual measurements.

The ‘Bland & Altman graphs’ also displayed, and the statistical analyses revealed, heteroscedasticity for the paretic limb in three of the four concentric measurements: the higher the test value the larger was the variability between the test sessions. Heteroscedasticity often occurs as larger measurement values by nature can give rise to larger absolute variability. If it occurs, it has been suggested that a logarithmic transformation of the data can be done before the analysis.[4, 19] However, Pearson's correlation coefficient was small for these three measurements, which imply that this heteroscedasticity may have little clinical relevance.

Even though an ICC value is high, it does not imply that a test is suitable for clinical use. Equipment and methods should display small measurement errors and should be sufficiently sensitive to detect real changes in muscle strength both for groups of subjects and single individuals. Several indices have been suggested for the evaluation of measurement errors. In the present study, we used the SEM and the SEM%, which gives the measurement errors in absolute and relative values. The SEM% represents the limit for the smallest change that indicates a real improvement for a group of subjects following, for example, an intervention. The SEM% values in this study were low for eccentric and concentric knee extension (between 9% and 12%), and somewhat higher for concentric knee flexion (between 14% and 20%). From a clinical standpoint, all these values seem reasonable and confirm that concentric knee muscle strength measurements – both flexion and extension – as well as eccentric knee extension strength can be used to detect real changes for a group of stroke subjects.

To detect a real change for a single subject, we calculated the smallest real difference (SRD).[20] The SRD% is independent of the unit of measurement and, like the SEM%, may be more easily interpreted. For the ten measurements evaluated in this study, the size of the relative change (SRD%) should exceed 22% (eccentric extension 60º/s, paretic limb) up to 55% (concentric flexion 120º/s, paretic limb) to indicate a real change for a single subject: if a subject performs 75 Nm with the paretic limb in concentric extension at 120º/s, this subject must improve 23 Nm to indicate a real change. From a clinical standpoint, only the SRD% values for the concentric and eccentric knee extension measurements are sufficiently small to be used to detect real changes in knee muscle strength for single subjects with mild to moderate post-stroke hemiparesis. The SRD% values for the knee flexion measurements may be too high to be useful for the same purpose. This clearly illustrates the need for several statistical indices to fully assess the reliability of a measurement method.

It has been reported that stroke patients cannot perform at higher angular velocities due to spastic antagonist restraints.[25] Studies of concentric knee muscle strength measurements in chronic stroke patients have therefore used low angular velocities (30-90º/s).[12-14] All subjects in the present study had a fairly good motor recovery and could perform the concentric contractions at the low velocity (60º/s), but a few subjects were unable to perform at the higher velocity (120º/s) and the eccentric contractions. Therefore, isokinetic dynamometry may be limited to fairly well recovered stroke subjects. Further studies are needed to assess the reliability of muscle strength measurements in stroke patients with a wider spectrum of post-stroke disability. Until then, our results should be restricted to fairly well-recovered post-stroke individuals. In addition, other muscle groups have to be tested, as reliability was only assessed for the knee muscles in the present study. Finally, the reliability of other aspects of the isokinetic strength measurement has to be established to fully address different components of functional muscle performance.

Acknowledgements

This study was supported by grants from the Swedish Stroke Association, Magn. Bergvall Foundation, the Swedish Association of Neurologically Disabled (NHR), the Swedish Society of Medicine, Gun and Bertil Stohne Foundation, the Crafoord Foundation, the Council for Medical Health Care Research in South Sweden, Lund University Hospital and Region Skåne.

Clinical message

In subjects with chronic mild to moderate post-stroke hemiparesis:

i) isokinetic knee muscle strength can be reliably measured at low to moderate angular velocities.

ii) isokinetic dynamometry can be used to detect real changes in knee muscle strength following interventions.

iii) only knee extension strength measurements are sufficiently sensitive to detect real changes in a single subject.

References

1 Patten C, Lexell J and Brown HE. Weakness and strength training in persons with post-stroke hemiplegia: Rationale, method and efficacy. *J Rehabil Res Develop* 2004; **41**: 293-312

2 Morris SL, Dodd KJ and Morris ME. Outcomes of progressive resistance strength training following stroke: A systematic review. *Clin Rehabil* 2004; **18**: 27-39

3 Rothstein JM. Measurement and clinical practice: Theory and application. In: Rothstein JM ed. *Measurement in physical therapy*. New York: Churchill Livingstone, 1985:1-46.

4 Atkinson G and Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998; **26**: 217-38

5 Bland JM and Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; **8**: 135-60

6 Flansbjer U-B, Holmbäck AM, Downham D, Patten C and Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. *J Rehabil Med* 2005;**37**:75-82

7 Holmbäck AM, Porter MM, Downham D and Lexell J. Reliability of isokinetic ankle dorsiflexor strength measurements in healthy young men and women. *Scand J Rehabil Med* 1999; **31**: 229-39

8 Holmbäck AM, Porter MM, Downham D and Lexell J. Ankle dorsiflexor muscle performance in healthy young men and women: Reliability of eccentric peak torque and work measurements. *J Rehabil Med* 2001; **33**: 90-96

9 Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000; **30**: 1-15

10 Rankin G and Stokes M. Reliability of assessment tools in rehabilitation: An illustration of appropriate statistical analyses. *Clin Rehabil* 1998; **12**: 187-99

11 Dvir Z. *Isokinetics: Muscle testing, interpretation, and clinical application.* New York: Churchill Livingstone, 1995

12 Eng JJ, Kim CM and Macintyre DL. Reliability of lower extremity strength measures in persons with chronic stroke. *Arch Phys Med Rehabil* 2002; **83**: 322-28

13 Hsu AL, Tang PF and Jan MH. Test-retest reliability of isokinetic muscle strength of the lower extremities in patients with stroke. *Arch Phys Med Rehabil* 2002; **83**: 1130-37

14 Pohl PS, Startzell JK, Duncan PW and Wallace D. Reliability of lower extremity isokinetic strength testing in adults with stroke. *Clin Rehabil* 2000; **14**: 601-07

15 Tripp EJ and Harris SR. Test-retest reliability of isokinetic knee extension and flexion torque measurements in persons with spastic hemiparesis. *Phys Ther* 1991; **71**: 390-96

16 Guide for the uniformed data set for medical rehabilitation FIM[SM]. (Swedish version 5.0). State University of New York at Buffalo.1996

17 Blackburn M, van Vliet P and Mockett SP. Reliability of measurements obtained with the Modified Ashworth Scale in the lower extremities of people with stroke. *Phys Ther* 2002; **82**: 25-34

18 Shrout PE and Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420-28

19 Bland M. *An introduction to medical statistics, third edition*. Oxford: Oxford University Press, 2000

20 Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD and Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res* 2001; **10**: 571-78

21 Fleiss JL. *The design and analysis of clinical experiments*. New York: John Wiley & Sons, 1986

22 Li RC, Wu Y, Maffulli N, Chan KM and Chan JL. Eccentric and concentric isokinetic knee flexion and extension: A reliability study using the Cybex 6000 dynamometer. *Br J Sports Med* 1996; **30**: 156-60

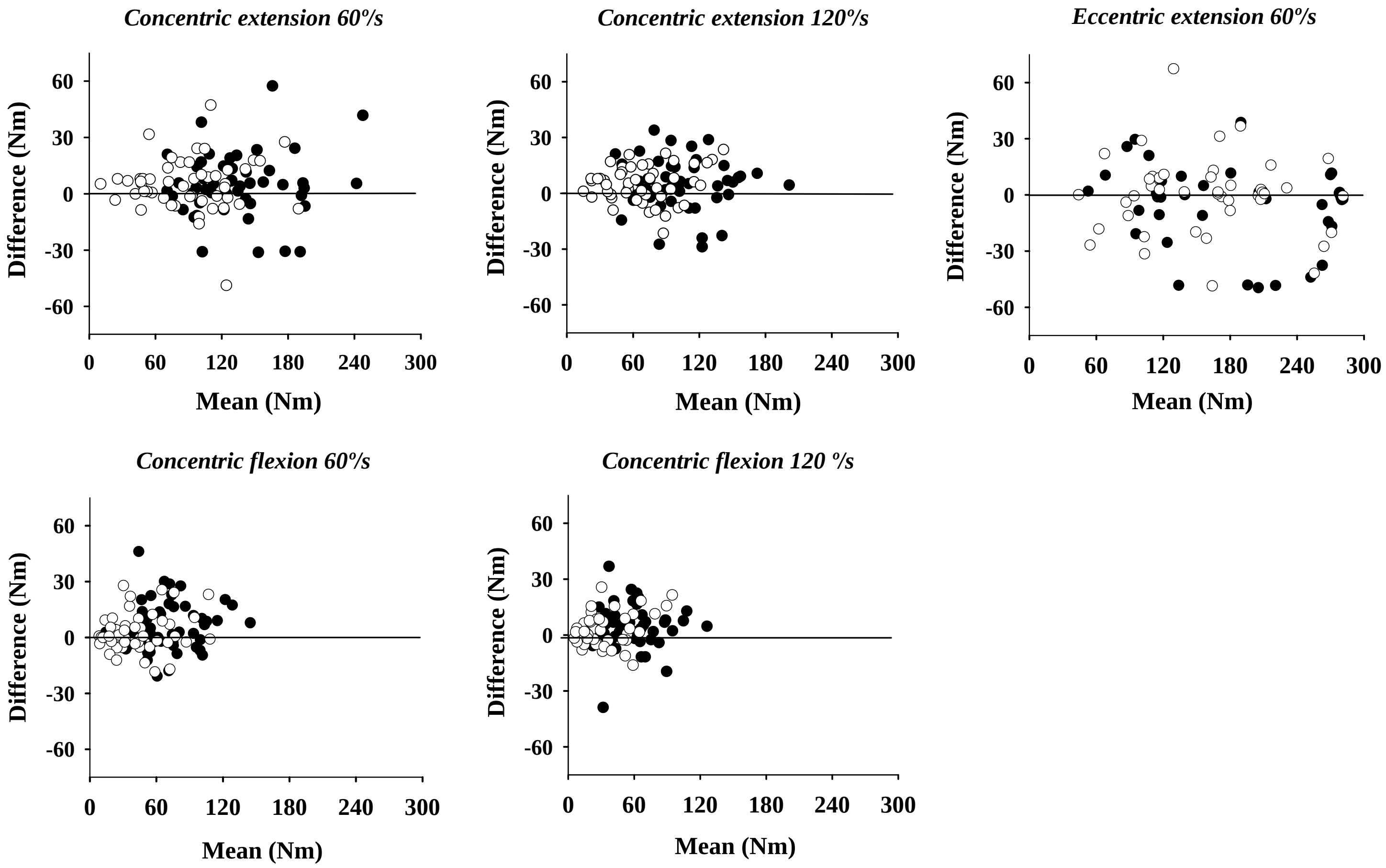
23 Pincivero DM, Lephart SM and Karunakara RA. Reliability and precision of isokinetic strength and muscular endurance for the quadriceps and hamstrings. *Int J Sports Med* 1997; **18**: 113-17

24 Steiner LA, Harris BA and Krebs DE. Reliability of eccentric isokinetic knee flexion and extension measurements. *Arch Phys Med Rehabil* 1993; **74**: 1327-35

25 Knutsson E, Martensson A and Gransberg L. Influences of muscle stretch reflexes on voluntary, velocity-controlled movements in spastic paraparesis. *Brain* 1997; **120** (Pt 9): 1621-33

Legends

Figure 1 The differences between test sessions 2 and 1 (test 2 minus test 1) plotted against the means of the two test sessions for the five different angular velocities for the non-paretic limb (filled circles) and the paretic limb (open circles). From these 'Bland & Altman graphs', the systematic variation around the zero line was revealed. There was evidence of heteroscedasticity, i.e. a significant larger variability for higher test values, from the paretic limb in three of the concentric angular velocities 60º/s extension (r=0.31), 120º/s extension (r=0.41) and 120º/s flexion (r=0.54).



Concentric extension 60°/s
Concentric extension 120°/s
Eccentric extension 60°/s
Concentric flexion 60°/s
Concentric flexion 120 °/s
Difference (Nm)
Mean (Nm)
60
30
0
-30
-60
0
60
120
180
240
300

Figure 1

Table 1 Reliability of the isokinetic concentric and eccentric knee muscle strength measurements for the 38 men and the 12 women

Measurement	$ICC_{2.1}$	95% CI for ICC	$\bar{d}$	95% CI for $\bar{d}$	SEM	SEM%	95% SRD	SRD%
Concentric knee extension at 60º/s								
Non-paretic	0.92	0.87 to 0.95	4.50	-0.38 to 9.38	12.4	9.3	-30.0 to 39.0	26
Paretic	0.94	0.89 to 0.96	5.06	1.00 to 9.12	10.6	11.9	-24.4 to 34.5	33
Concentric knee extension at 120º/s								
Non-paretic	0.93	0.87 to 0.96	4.29	0.44 to 8.14	10.0	9.8	-23.3 to 31.9	27
Paretic [a]	0.95	0.92 to 0.97	5.00	2.15 to 7.85	7.7	11.2	-16.4 to 26.4	31
Concentric knee flexion at 60º/s								
Non-paretic	0.89	0.81 to 0.93	5.91	2.17 to 9.65	10.1	14.2	-22.1 to 34.0	39
Paretic	0.92	0.87 to 0.96	2.52	-0.38 to 5.42	7.4	17.4	-17.9 to 22.9	48
Concentric knee flexion at 120º/s								
Non-paretic	0.89	0.82 to 0.94	4.00	0.67 to 7.34	8.7	15.3	-20.1 to 28.1	42
Paretic [a]	0.93	0.88 to 0.96	2.50	-0.01 to 5.01	6.3	20.0	-15.0 to 20.0	55
Eccentric knee extension at 60º/s								
Non-paretic [b]	0.96	0.93 to 0.98	-4.77	-10.97 to 1.43	14.7	7.8	-45.4 to 35.8	22
Paretic [b]	0.95	0.90 to 0.97	-0.12	-6.43 to 6.19	14.5	9.1	-40.3 to 40.1	25

[a] n=48

[b] n=44

ICC, intraclass correlation coefficient; 95% CI, 95% confidence interval; SEM, standard error of measurement; SRD, smallest real difference