Reproducibility and validity of measuring the length of upper extremity flexor muscles of the spastic forearm in acquired brain injury patients

Masterthesis

Physiotherapy Science

Program in Clinical Health Sciences

Utrecht University

Name student:	E.H. Koolen, Bsc.
Student number:	3848442
Date:	27-06-2014
Internship supervisor(s):	Dr. I. van de Port
Internship institute:	Revant Revalidatiecentrum Breda, the Netherlands
Lecturer/supervisor Utrecht University:	Dr. M.F. Pisters

"ONDERGETEKENDE

Noortje Koolen,

bevestigt hierbij dat de onderhavige verhandeling mag worden geraadpleegd en vrij mag worden gefotokopieerd. Bij het citeren moet steeds de titel en de auteur van de verhandeling worden vermeld."

Examiner

Dr. M.F. Pisters

Assessors:

Dr. I. van de Port

Dr. J. van der Net

Masterthesis, Physiotherapy Science, Program in Clinical Health Sciences, Utrecht University, Utrecht, 2014

SAMENVATTING

Doelstelling: Het doel is om de reproduceerbaarheid en criterium validiteit te bepalen van een nieuw meetinstrument dat de maximale passieve extensie van de pols meet, bij patiënten met spasme in de spieren van de onderarm na niet-aangeboren hersenletsel.

Methode: De metingen met het nieuwe instrument werden verricht door vier beoordelaars. Om de interbeoordelaars betrouwbaarheid te bepalen werd de patiënt tweemaal gemeten door twee verschillende beoordelaars binnen maximaal 24 uur. De intrabeoordelaars betrouwbaarheid werd bepaald door een van de vier beoordelaars binnen maximaal 1 week. De resultaten van de eerste metingen werden vergeleken met de metingen uitgevoerd door de behandeld fysiotherapeut met de manuele goniometer. De statistiek bestond uit de Intraclass Correlation Coefficients (ICC), Limits of Agreement (LoA), Standard Error of Measurements (SEM) en Smallest Detectable Changes (SDC).

Resultaten: ICCs werden beoordeeld als uitstekend voor zowel interbeoordelaars (ICC [2.1]: 0.933) als intrabeoordelaars betrouwbaarheid (ICC [3.1]: 0.990). De interbeoordelaars overeenstemming resulteerde in LoA van -19.3° tot 20.4°, SEM van 7.3° en SDC van 20.2°. Intrabeoordelaars overeenstemming resulteerde in LoA van -7.5° tot 7.5°, SEM van 2.7° en SDC van 7.5°. De criterium validiteit was goed (Spearman rank correlation coefficient, r_s: 0.772)

Conclusie: De studie resulteerde in een goede reproduceerbaarheid en criterium validiteit van het nieuwe meetinstrument. Echter, om een duidelijkere uitspraak te kunnen doen over reproduceerbaarheid en validiteit is het noodzakelijk meer patiënten te includeren in toekomstige studies.

Klinische relevantie: Met dit nieuwe betrouwbare en valide meetinstrument kan de lengte van de flexoren van de onderarm bij patiënten met spasme gemeten worden door één persoon, in plaats van dat deze meting door twee personen uitgevoerd moet worden.

ABSTRACT

Aim: The aim of this study is to determine the reproducibility and the criterion validity of a new instrument which measures the length of flexor muscles of the spastic forearm in acquired brain injury patients.

Methods: To determine the inter-rater reproducibility four raters measured with this new instrument. After the first measurement by the first rater, a second rater measured the same participant after at least 30 minutes and maximally 24 hours. One of the four raters, this was always the same rater, repeated the measurements within one week to determine the intra-rater reproducibility. The criterion validity was determined by a measurement performed by the regular physiotherapist with a manual goniometer. Statistical analyses were done by evaluating the Intraclass Correlation Coefficients (ICC), Limits of Agreement (LoA), Standard Error of Measurement (SEM) and Smallest Detectable Change (SDC).

Results: ICCs revealed an excellent inter-rater (ICC [2.1]: 0.933) and intra-rater reliability (ICC [3.1]: 0.990). Inter-rater agreement resulted in LoA of -19.3°to 20.4°, SEM of 7.3° and SDC of 20.2°. Intra-rater agreement revealed LoA of -7.5° to 7.5°, SEM of 2.7° and SDC of 7.5°. Criterion validity with the manual goniometer was indicated as good (Spearman rank correlation coefficient, r_s: 0.772).

Conclusion: This study confirms an overall good reproducibility and criterion validity of a new instrument in a population in brain injury patients with a spastic wrist. However, a larger sample size is necessary to draw firm conclusions about the reproducibility and validity of this instrument.

Clinical Relevance: Besides the measurements properties, the clinical relevance of this study implies that this new instrument has the potential to be used in clinical practice by only one health care professional to determine the muscle length of the spastic wrist.

Keywords: *measurement properties, spastic forearm, range of motion.*

INTRODUCTION

Acquired brain injury is a worldwide problem and one of the leading contributors to adult disability and death. It results from various causes including traumatic brain injury, stroke, cerebral tumour, cerebral anoxia and encephalitis (1). The survivors commonly experience long-term impairments that affect their quality of life (2-6). Spasticity can be one of these long-term impairments and is common in the upper extremity, especially after stroke (7-9). The current definition of spasticity is described by Pandyan et al as 'a phenomenon of disturbed sensory-motor control of muscle tone connected with upper motor neuron damage resulting in intermittent or sustained, involuntary, muscle tone hyperactivity' (10). The prevalence of post-stroke spasticity after a first stroke ranges from 18% to 42.4%, of which 4% suffers from disabling spasticity (7,11-16). Eventually, a prolonged immobilization of a joint due to spasticity may cause contractures (17). Wrist flexion contractures can develop rapidly within 6–8 weeks (18)(10).

Spasticity and contractures of the upper limb often result in difficulties in daily activities (19). A permanent shortening of the flexor muscles of the wrist and fingers can lead to a clenched fist (20,21), which causes pain and may result in poor hygiene of the palmar skin of the hand (20). Therefore during the rehabilitation process it is important to prevent contractures of the upper limb and to control spasticity.

To prevent contractures or treat spasticity of the forearm during the rehabilitation process it is important to measure the changes in wrist motion by evaluating the passive extension degrees of the wrist. Nowadays, in daily practice the wrist motion is measured with a manual goniometer (22,23), but this is difficult in patients with a spastic forearm. Two persons are required to measure the maximum passive range of motion. One person should stretch the wrist and fixate the hand and the other person should place the goniometer and read the degrees to determine the passive extension degrees of the wrist.

A new instrument to measure the length of the flexors of the forearm has been developed by Revant Rehabilitation centre Breda in cooperation with Orthopedisch Instrument Makerij (OIM) Brabant Breda, see Appendix A. This new instrument consists of a goniometer, which must be fixated to the forearm. The maximum possible passive extension of the wrist should be determined when fingers and hand are completely stretched to evaluate the muscle length of the forearm. With this new instrument only one health care professional is needed to determine the muscle length. Experiences from a pilot study and regular care showed that measuring with this method was not painful and was well tolerated by the patient.

Therefore, since the usability of this measurement has already been established, the reproducibility and validity need further investigation. The aim of this study was to determine the

reproducibility and the criterion validity of a new instrument which measures the length of flexor muscles of the spastic forearm in acquired brain injury patients who are treated at Revant Rehabilitation centre Breda or Lindenhof Rehabilitation centre Goes.

METHODS

This study is a quantitative reproducibility study to demonstrate reliability, agreement and validity, with a test re-test design, of a new developed measuring instrument. This new instrument has been developed by Revant Rehabilitation centre Breda in cooperation with OIM Brabant Breda, see Appendix A. The ethical committees of Maxima Medisch Centrum, Eindhoven-Veldhoven, the Netherlands, concluded that the study was not subject to the Medical Research Involving Human Subjects Act (WMO).

A convenience sample of patients after brain injury was used within this study. In order to be eligible to participate, a subject had to meet the following criteria: 1. Survived an acquired brain injury, 2. Time since acquired brain injury was more than one month, 3. Received in- or outpatient rehabilitation at Revant Rehabilitation centre Breda or Lindenhof Rehabilitation centre Goes, 4. Age of 18 years or older, 5. Spasticity in the forearm, which was measured with the Modified Modified Ashworth Scale (MMAS) by an experienced physiotherapist and the score had to be at least 1. Patients were excluded if: 1. They had wounds in the spastic forearm, 2. Restrictions in mobility of the affected forearm or joint pathologies other than directly related to spasticity, 3. Communication difficulties or language barriers and 4. Patients received a botulinum toxin injection or surgery as treatment for spasticity.

Study parameters
Main study parameter was the maximum passive extension of the spastic wrist measured with the new instrument to determine the inter- and intra-rater reliability and agreement.
To determine the criterion validity the maximum passive extension in degrees of the wrist was also measured with a manual goniometer. Measuring with a manual goniometer is the current regular care. The values were noted with an accuracy of one degree.
The secondary parameters were the Modified Modified Ashworth Scale (MMAS) (24-27), the Motricity Index (MI) of the upper extremity (28), and an overall evaluation of pain at each end range of motion measured with this new instrument (score 0 if there was no pain and score 1 if there was pain). Other study parameters included the characteristics of participants, ie. age, gender, time since brain injury, diagnosis, affected side and rehabilitation centre.

Study procedure
The measurements with this new instrument were done during regular therapy in a physical therapy treatment room at Revant Rehabilitation centre Breda and at Lindenhof Rehabilitation centre Goes. Four independent physiotherapists performed the measurements (NK, AA, AB, PW). The raters practiced in advance to familiarize with and discuss the method.

To determine the inter-rater reliability and agreement, two raters measured the maximum passive extension degrees. After the first measurement was performed by the first rater, the second rater measured the same participant after at least 30 minutes and maximally 24 hours. The two raters were blinded for each other's results.

To determine the intra-rater reliability and agreement one rater, repeated the measurement within one week. Intra-rater parameters were determined by the same rater for all participants (NK).

Before a measurement took place, the rater started with a 'warming-up' to mobilize the upper extremity of the participant. The upper extremity was passively moved for at least two minutes in order to reduce the spasticity. The patient was sitting in a (wheel)chair without armrests. The shoulder was in neutral position and the elbow in 90 degrees of flexion. Distal interphalangeal-, proximal interphalangeal-, and metacarpophalangeal joints were in neutral positions. The instrument was fixated on the forearm, wrist and fingers, in which the fingers were fully stretched, see Appendix A. One hand of the rater fixated the proximal part of the instrument on the forearm of the patient. The instrument itself was fixated beyond the processus styloideus ulnae. The other hand of the rater fixated the distal part on the forearm and moved fingers toward dorsal flexion. Each measurement took about five to ten minutes. During each measurement the rater determined the passive extension of the spastic wrist three times. The final outcome was the mean of three range of motions expressed in degrees.

Besides reliability measurements, the physiotherapists who actually treated the patients suffering from brain injury at Revant Rehabilitation centre Breda or Lindenhof Rehabilitation centre Goes measured the maximum passive extension of the wrist to determine the criterion validity. They measured three times the range of motion with a manual goniometer as done in usual care. They reported their results on a form in the research file and were blinded for the results that were obtained with the new instrument by the other raters.

Statistical analysis:

Descriptive statistics were used to describe the characteristics of the study population.

This study examined the reproducibility, based on inter- and intra-rater reliability and agreement, of this new instrument. Reliability and agreement are two different parameters. Reliability parameters focus on how well patients can be distinguished from each other, despite measurement error (29). And agreement parameters focus on measurement error and asses exactly how close the scores for repeated measurements are (29). As an umbrella term for the concepts of reliability and agreement the term 'reproducibility' was used within this study, because both concepts concern the question of whether measurement results are reproducible in test-retest situations (29).

The inter- and intra-rater reliability were calculated with Intraclass Correlation Coefficients (ICC) on data that have been produced from a two-way analysis of variance (ANOVA) (30). See Appendix B for a description of the statistical formulas. ICC [2.1] has been calculated to determine the inter-rater reliability, because each participant has been rated by four raters who were randomly chosen and therefore the results can be generalized to other raters with similar characteristics (30-32). An ICC [3.1] has been calculated to determine the intra-rater reliability, because this type of ICC ignores systematic differences between the raters (29,31,32). ICC values were classified as suggested by Portney and Watkins (32), excellent for ICC ≥ 0.90, good for ICC ≥ 0.75, moderate for ICC ≥ 0.5 and poor for ICC < 0.5.

An ICC alone gives no indication of the magnitude of disagreement between raters, therefore inter- and intra-rater agreement were calculated based on Bland and Altman 95% limits of agreement tests, a calculation of the Standard Error of Measurements (SEM) and Smallest Detectable Changes (SDC) were complemented (29,33). Bland-Altman Plots were made for all measurements by using the Limits of Agreement (LoA) of 1.96 standard deviations (SD) above and below the mean difference (33). $SEM_{agreement}$ was calculated on the data of the inter-rater reliability and $SEM_{consistency}$ was calculated based on the data of the intra-rater reliability, see Appendix B (29). The SDCs were also calculated, based on the calculations of $SEM_{agreement}$ and $SEM_{consistency,}$, see Appendix B (29).

Besides the reproducibility, the criterion validity was examined within this study. Criterion validity analysis was performed to determine if both measures, our new instrument and manual goniometer, produced comparable results. Therefore the correlation coefficient between the passive extension in degrees measured with the new instrument and the manual goniometer was calculated. A scatter plot has been created to visually clarify the strength and shape of the relationship. The Spearman rank correlation coefficient, given the symbol r_s (Spearman's rho), has been determined because of the small sample size within this study. Correlation coefficients can range from -1.00 (perfect negative relationship), to 0.00 (no correlation), to +1.00 (perfect positive relationship). The correlation has been classified: 0.00-0.25; little or no relationship, 0.25-0.50; fair relationship, 0.50-0.75; moderate to good relationship and above 0.75; good to excellent relationship (32).

All statistics were done using Statistical Package for the Social Solutions software (SPSS Statistics, version 21.0, IBM corporation, Somers USA).

RESULTS

Details of 12 participants were available for analysis. Eleven patients were diagnosed with stroke (34) and 1 patient survived a brain tumour. Of these 12 patients, 6 (50%) were female with a median age 59.5 years (Interquartile Range (IQR): 16). Nine (75%) of them received in- or outpatient rehabilitation at Revant Rehabilitation centre Breda, 3 (25%) at Lindenhof Rehabilitation centre Goes. Nine patients (75%) were affected on the right side of the body. Overall, the patients within this study obtained a low score on the MI. Based on the results of the MMAS, 4 patients (33.3%) obtained a score of 1 for wrist spasticity, 7 (58.3%) of them obtained a score of 2 and only 1 patient (8.3%) a score of 3. The evaluation of pain obtained during the measurements showed that most of the patients gave the same judgement of pain during all 3 testing moments. Descriptive characteristics of the participants and the secondary parameters are summarized in Table 1.

Reliability
Inter-rater reliability was based on data of 12 patients and the intra-rater reliability was based on 11 patients. One patient was already discharged from the rehabilitation centre at the time the second measurement should took place. The ICCs are presented in Table 2.
The inter-rater reliability was classified as excellent (ICC values ≥ 0.90) with an ICC [2.1] of 0.933 (95% confidence interval (CI): 0.783-0.980). The intra-rater reliability was also classified as excellent, with an ICC [3.1] of 0.990 (95% CI: 0.963-0.997).

Agreement
The 95% LoA, $SEM_{agreement}$ and $SEM_{consistency}$ and the values of SDCs, are also presented in Table 2. The inter-rater agreement was based on 95% LoA, presented in a Bland and Altman plot (fig.1), were between -19.3° and 20.4°. This means by definition that approximately 95% of the differences between the repeated measurements, obtained by four independent raters, with this new instrument lie within -19.3° and 20.4°. A $SEM_{agreement}$ of 7.3° was calculated. Based on this SEM a SDC value of 20.2° was found.
The intra-rater agreement, based on LoA, were between -7.5° and 7.4° and are also presented in a Bland and Altman plot (fig.2). A $SEM_{consistency}$ value of 2.7° was calculated together with a SDC value of 7.5°.

Criterion validity
Criterion validity was based on data of 12 patients. Table 3 shows that the range of motions of the wrist obtained with the new instrument and the manual goniometer were positively correlated (r_s=0.772, p=0.003). A value of r_s 0.772 was indicated as a good significant relationship, visualized by a scatter plot, see fig.3.

DISCUSSION

The aim of this study was to determine the reproducibility, based on inter- and intra-rater reliability and agreement, and the criterion validity of a new instrument that measures the length of flexor muscles of the spastic forearm in patients suffering from acquired brain injury. Based on our findings, the conclusion can be justified that this new instrument has an excellent inter- and intra-rater reliability, both values of ICC were above or equal 0.90. Many studies have demonstrated that intra-tester reliability within goniometric measurements is most of the time higher than inter-tester reliability, which is also an outcome within this study (22). The inter-tester reliability shows greater inter-tester variance than intra-tester variance.

The Bland and Altman plots for inter- and intra-rater agreement, showed that one person obtained twice an average of less than 0 and all other measurements had an average greater than 0, see fig.1 and fig.2. The averages lower than 0 were caused by the degree of spasticity in the forearm by this one patient and therefore negative passive range of motions of the wrist were obtained. The degree of spasticity was evaluated with a MMAS of 3. Only one person obtained a MMAS of 3 within this study. In future studies more patients with a MMAS of $\geq$ 3 should be included to ensure a homogeneous group of patients and to make sure that not only one person provides a certain effect on the statistics.

The SEM varied between de inter- and intra-rater agreement. A higher score was obtained for the inter-rater agreement ($SEM_{agreement}$: 7.3°) compared with the intra-rater agreement ($SEM_{consistency}$: 2.7°). The $SEM_{agreement}$ takes the systematic differences between the 4 raters into account and $SEM_{consistency}$ ignores this systematic difference (29). These SEMs are considered acceptable based on a previous study that determined the maximum passive wrist motion within healthy probands (35). The SDC can be explained as a value that specifies with 95% certainty that the observed change is not caused by a random measurement error but as a true change in range of motion. Because of a higher value of $SEM_{agreement}$ compared with $SEM_{consistency}$, also the SDC value for the inter-rater agreement was higher (SDC: 20.2°) compared with the SDC for the intra-rater reliability. This can be due to the small sample size within this study and the lack of experience with this new instrument. It could have been helpful to practice more in advance to familiarize with the method.

Furthermore, the measurements with this new instrument and measurements with the manual goniometer revealed a good positive significant relationship, based on a Spearman correlation coefficient of 0.772 and a significance of 0.003. Within this study the measurements with this new instrument were assessed by four different raters and the measurements with the manual goniometer were assessed by six different raters. Although a higher correlation would have been obtained when the measurements with our new instrument obtained by the same rater were correlated with the measurements with the manual goniometer, the correlation coefficient of our study shows the results of clinical practice.

Spasticity is frequently assessed by subjectively evaluating a catch and release or by estimating the resistance at the range of motion of the spastic joint. The spasticity is then evaluated on a subjective scale, such as the (Modified) Tardieu Scale or the (Modified Modified) Ashworth Scale. The reliability and validity of these subjective scales are questionable (36-38). Besides these subjective scores obtained to assess spasticity, it is also important to objectify the range of motion of the spastic joint. Therefore, to prevent contractures or treat spasticity of the forearm during the rehabilitation process it is necessary to objectify the changes in mobility with evaluating the passive range of motion in degrees of the spastic joint, which can be done with this new instrument.

There is, to our knowledge, only one study that evaluated repeated measurements with a hydrogoniometer of arm joint passive range of motions in stroke patients with a paretic arm (39). The arm joint passive range of motions also consisted of the passive wrist extension with extended fingers, as in our study. This study by De Jong et al (39) found an inter-rater ICC of 0.93 (95% CI: 0.90-0.96). Within our study also an inter-rater ICC of 0.933 was obtained with our new instrument, but with a 95% CI of 0.789-0.980. A narrow interval gives a better estimate of the true degree of reliability than a wide interval. This can also be due to the small sample size. Small sample sizes will yield an imprecise estimate of the reliability coefficients which is indicated by a wide confidence interval (CI)(40). Furthermore they found an overall SEM of 4.6° and a SDC of 12.8°. These values are lower than the results within this study (SEM$_{agreement}$: 7.3° and SDC: 20.2°). Besides the difference in sample sizes there were other differences between these two studies. Within the study by De Jong et al (39) the measurements were done with a hydrogoniometer, the patients were not specifically included based on the presence of a spastic forearm and the measurements were assessed by two raters. One rater passively moved the wrist to extension and the other rater observed the degrees of range of motion. Within our study, only one rater assessed the passive wrist extension in patients with a spastic forearm, this person moved the wrist to extension and observed the degrees of range of motion, and that seems to be just as reliable.

Besides the comparisons of our results with the study by De Jong et al (39), there are other studies in which the reliability of the manual goniometer, which measures the mobility of the wrist, has been determined. However, the difference between these studies is that patients without spasm were included. One study determined the inter- and intra-rater SEM of the manual goniometer based on wrist motion in healthy subjects (35). An inter-rater SEM of 5° up to 8° and an intra-rater SEM of 6° up to 10° were found (35). Within the study by LaStayo et al (41), based on patients with wrist complaints, radial, ulnar and volar alignment techniques with the manual goniometer when measuring passive wrist extension have been evaluated. The same inter-rater and intra-rater ICCs were obtained, for the radial alignment and ulnar alignment technique an ICC of 0.80, and an ICC of 0.84 for volar alignment technique. Furthermore, inter-rater SEMs of 7.69°, 7.87° and 6.00° were found and intra-rater SEMs of 7.82°, 8.07°, 5.57° were

obtained for the three techniques. Carter et al (23) evaluated also those three techniques when measuring the maximum passive wrist extension, respectively the radial, ulnar and dorsal/volar alignment techniques. They obtained inter-rater ICCs of 0.3, 0.5 and 0.9, and different intra-rater ICCs of 0.8 up to 1 were obtained. Furthermore they found root mean squared error (RMSE) of 7° up to 8°. The results by Carter et al (23), were based on cadaveric upper extremities. Within our study, similar or even higher results of reproducibility were obtained compared with these three studies, while our results are even based on patients with a spastic forearm. Especially if our results were compared with the results obtained with the ulnar alignment techniques, because this was the same technique that have been used within our study.

There are some limitation within this study. When interpreting the results from this study, it should be considered that there was a small sample size of patients included, which prohibits us from drawing firm conclusions. Within this study the standard method, a measurement with a manual goniometer, was chosen to analyze the criterion validity. However, the use of this method was not analyzed objectively. Based on the literature, the volar/dorsal alignment technique is the goniometric technique of choice (23,41). Despite that the manual goniometer is the known standard in clinical practice of measuring the maximum passive range of motion of the wrist, a clear explanation of the method would be helpful within future studies. Another limitation is that the four raters were experienced physiotherapists, this may have led to more positive results. However, also in clinical practice different levels of experience are often present. Although the power of our study is limited by the small number of patients, the design of our study closely reflects regular clinical practice.
More research is needed with a larger sample size to evaluate the reproducibility and validity of this new instrument in clinical practice in patients after acquired brain injury, especially after stroke, with a spastic forearm.

CONCLUSION

This study confirms an overall good reproducibility, based on inter- and intra-rater reliability and agreement, and criterion validity of a new instrument to measure the length of flexor muscles of the spastic forearm in acquired brain injury, and especially in stroke patients. Besides the measurements properties, the clinical relevance of this study implies that this new instrument has the potential to be used in clinical practice by only one health care professional to determine the muscle length of the spastic wrist.

REFERENCES

(1) Chevignard M, Toure H, Brugel DG, Poirier J, Laurent-Vannier A. A comprehensive model of care for rehabilitation of children with acquired brain injuries. Child Care Health Dev 2010 Jan;36(1):31-43.

(2) Murray J, Ashworth R, Forster A, Young J. Developing a primary care-based stroke service: a review of the qualitative literature. Br J Gen Pract 2003 Feb;53(487):137-142.

(3) Haacke C, Althaus A, Spottke A, Siebert U, Back T, Dodel R. Long-term outcome after stroke: evaluating health-related quality of life using utility measurements. Stroke 2006 Jan;37(1):193-198.

(4) Hare R, Rogers H, Lester H, McManus R, Mant J. What do stroke patients and their carers want from community services? Fam Pract 2006 Feb;23(1):131-136.

(5) Brown AW, Elovic EP, Kothari S, Flanagan SR, Kwasnica C. Congenital and acquired brain injury. 1. Epidemiology, pathophysiology, prognostication, innovative treatments, and prevention. Arch Phys Med Rehabil 2008 Mar;89(3 Suppl 1):S3-8.

(6) Dawson DR, Chipman M. The disablement experienced by traumatically brain-injured adults living in the community. Brain Inj 1995 May-Jun;9(4):339-353.

(7) Lundstrom E, Terent A, Borg J. Prevalence of disabling spasticity 1 year after first-ever stroke. Eur J Neurol 2008 Jun;15(6):533-539.

(8) Brainin M, Norrving B, Sunnerhagen KS, Goldstein LB, Cramer SC, Donnan GA, et al. Poststroke chronic disease management: towards improved identification and interventions for poststroke spasticity-related complications. Int J Stroke 2011 Feb;6(1):42-46.

(9) Graham LA. Management of spasticity revisited. Age Ageing 2013 Jul;42(4):435-441.

(10) Pandyan AD, Gregoric M, Barnes MP, Wood D, Van Wijck F, Burridge J, et al. Spasticity: clinical perceptions, neurological realities and meaningful measurement. Disabil Rehabil 2005 Jan 7-21;27(1-2):2-6.

(11) Sommerfeld DK, Eek EU, Svensson AK, Holmqvist LW, von Arbin MH. Spasticity after stroke: its occurrence and association with motor impairments and activity limitations. Stroke 2004 Jan;35(1):134-139.

(12) Leathley MJ, Gregson JM, Moore AP, Smith TL, Sharma AK, Watkins CL. Predicting spasticity after stroke in those surviving to 12 months. Clin Rehabil 2004 Jun;18(4):438-443.

(13) Watkins CL, Leathley MJ, Gregson JM, Moore AP, Smith TL, Sharma AK. Prevalence of spasticity post stroke. Clin Rehabil 2002 Aug;16(5):515-522.

(14) Welmer AK, von Arbin M, Widen Holmqvist L, Sommerfeld DK. Spasticity and its association with functioning and health-related quality of life 18 months after stroke. Cerebrovasc Dis 2006;21(4):247-253.

(15) Moura Rde C, Fukujima MM, Aguiar AS, Fontes SV, Dauar RF, Prado GF. Predictive factors for spasticity among ischemic stroke patients. Arq Neuropsiquiatr 2009 Dec;67(4):1029-1036.

(16) Ryu JS, Lee JW, Lee SI, Chun MH. Factors predictive of spasticity and their effects on motor recovery and functional outcomes in stroke patients. Top Stroke Rehabil 2010 Sep-Oct;17(5):380-388.

(17) Botte MJ, Nickel VL, Akeson WH. Spasticity and contracture. Physiologic aspects of formation. Clin Orthop Relat Res 1988 Aug;(233)(233):7-18.

(18) Pandyan AD, Cameron M, Powell J, Stott DJ, Granat MH. Contractures in the post-stroke wrist: a pilot study of its time course of development and its association with upper limb recovery. Clin Rehabil 2003 Feb;17(1):88-95.

(19) Schinwelski M, Slawek J. Prevalence of spasticity following stroke and its impact on quality of life with emphasis on disability in activities of daily living. Systematic review. Neurol Neurochir Pol 2010 Jul-Aug;44(4):404-411.

(20) Heijnen IC, Franken RJ, Bevaart BJ, Meijer JW. Long-term outcome of superficialis-to-profundus tendon transfer in patients with clenched fist due to spastic hemiplegia. Disabil Rehabil 2008;30(9):675-678.

(21) Petrella L, Belkheyar Z, Oberlin C. The psychoflexed hand: New perspectives in diagnosis, classification and treatment. Chir Main 2013 Sep;32(4):245-250.

(22) Gajdosik RL, Bohannon RW. Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. Phys Ther 1987 Dec;67(12):1867-1872.

(23) Carter TI, Pansy B, Wolff AL, Hillstrom HJ, Backus SI, Lenhoff M, et al. Accuracy and reliability of three different techniques for manual goniometry for wrist motion: a cadaveric study. J Hand Surg Am 2009 Oct;34(8):1422-1428.

(24) Ansari NN, Naghdi S, Hasson S, Fakhari Z, Mashayekhi M, Herasi M. Assessing the reliability of the Modified Modified Ashworth Scale between two physiotherapists in adult patients with hemiplegia. NeuroRehabilitation 2009;25(4):235-240.

(25) Ansari NN, Naghdi S, Mashayekhi M, Hasson S, Fakhari Z, Jalaie S. Intra-rater reliability of the Modified Modified Ashworth Scale (MMAS) in the assessment of upper-limb muscle spasticity. NeuroRehabilitation 2012;31(2):215-222.

(26) Naghdi S, Ansari NN, Azarnia S, Kazemnejad A. Interrater reliability of the Modified Modified Ashworth Scale (MMAS) for patients with wrist flexor muscle spasticity. Physiother Theory Pract 2008 Sep-Oct;24(5):372-379.

(27) Naghdi S, Ansari NN, Mansouri K, Asgari A, Olyaei GR, Kazemnejad A. Neurophysiological examination of the Modified Modified Ashworth Scale (MMAS) in patients with wrist flexor spasticity after stroke. Electromyogr Clin Neurophysiol 2008 Jan-Feb;48(1):35-41.

(28) Collin C, Wade D. Assessing motor impairment after stroke: a pilot reliability study. J Neurol Neurosurg Psychiatry 1990 Jul;53(7):576-579.

(29) de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006 Oct;59(10):1033-1039.

(30) Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil 1998 Jun;12(3):187-199.

(31) Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428.

(32) Portney LG, Watkins MP. Foundations of Clinical Research: applications to practice. Third edition ed. Upper Saddle River, New Jersey: Pearson Education; 2009.

(33) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310.

(34) Aho K, Harmsen P, Hatano S, Marquardsen J, Smirnov VE, Strasser T. Cerebrovascular disease in the community: results of a WHO collaborative study. Bull World Health Organ 1980;58(1):113-130.

(35) Solgaard S, Carlsen A, Kramhoft M, Petersen VS. Reproducibility of goniometry of the wrist. Scand J Rehabil Med 1986;18(1):5-7.

(36) Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther 1987 Feb;67(2):206-207.

(37) Ansari NN, Naghdi S, Moammeri H, Jalaie S. Ashworth Scales are unreliable for the assessment of muscle spasticity. Physiother Theory Pract 2006 Jun;22(3):119-125.

(38) Naghdi S, Ansari NN, Abolhasani H, Mansouri K, Ghotbi N, Hasson S. Electrophysiological evaluation of the Modified Tardieu Scale (MTS) in assessing poststroke wrist flexor spasticity. NeuroRehabilitation 2014;34(1):177-184.

(39) de Jong LD, Dijkstra PU, Stewart RE, Postema K. Repeated measurements of arm joint passive range of motion after stroke: interobserver reliability and sources of variation. Phys Ther 2012 Aug;92(8):1027-1035.

(40) Eliasziw M, Young SL, Woodbury MG, Fryday-Field K. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. Phys Ther 1994 Aug;74(8):777-788.

(41) LaStayo PC, Wheeler DL. Reliability of passive wrist flexion and extension goniometric measurements: a multicenter study. Phys Ther 1994 Feb;74(2):162-74; discussion 174-6.

Tabel 1: Characteristics of the included patients

Characteristics (n=12)			
	Median	(IQR)[a]	
Age, (y)	59.5	(16)	
Diagnosis, n(%)			
Stroke	11 (91.7)		
Brain tumor	1 (8.3)		
Gender, n(%)			
Female	6 (50)		
Male	6 (50)		
Affected Side, n(%)			
Left	3 (25)		
Right	9 (75)		
Rehabilitation centre, n(%)			
Breda	9 (75)		
Goes	3 (25)		
MI[b], pinch grip	5.5	(11)	
MI, elbow flexion	9.0	(5)	
MI, shoulder abduction	11.5	(5)	
MI total	28.5	(20)	
MMAS[c], n(%)	Elbow	Wrist	Finger
0	1 (8.3)	0	0
1	6 (50)	4 (33.3)	6 (50)
2	4 (33.3)	7 (58.3)	4 (33.3)
3	1 (8.3)	1 (8.3)	2 (16.7)
4	0	0	0
Evaluation of pain, n(%)	Test 1	Test 2	Test 3
Yes	5 (41.7)	5 (41.7)	4 (33.3)
No	7 (58.3)	7 (58.3)	7 (58.3)
Missing			1 (8.3)

[a] IQR = Interquartile Range, [b] MI = Motricity Index

[c] MMAS = Modified Modified Ashworth Scale. 0 = no increase in muscle tone, 1 = slight increase in muscle tone, manifested by a catch and release or by minimal resistance at the end range of motion when the affected part(s) is moved in flexion or extension, 2 = marked increase in tone, manifested by a catch in the middle range and resistance throughout the remainder of the range of motion, but affected part(s) easily moved, 3 = considerable increase in tone, passive movement difficult, 4 = affected part(s) rigid in flexion or extension

Tabel 2. Inter- and intra-rater reliability and agreement based on results of the measurements with the new instrument

	Inter-rater (n=12)	Intra-rater (n=11)
Reliability		
ICC[a] (95% CI)[b]	0.933 (0.783-0.980)	0.990 (0.963-0.997)
Agreement		
LoA[c]	- 19.3° – 20.4°	- 7.5° – 7.4°
SEM[d]	7.3°	2.7°
SDC[e]	20.2°	7.5°

[a] ICC = Intraclass Correlation Coefficient, [b] (95% CI) = 95% Confidence Interval, [c] LoA = Limits of Agreement: 1.96 standard deviations (SD) above and below the mean difference, [d] SEM = Standard Error of Measurement, [e] SDC = Smallest Detectable Change

Tabel 3: Criterion validity, based on the results of the measurements with the new instrument correlated with the measurements with the goniometer

|	New instrument Test 1		Goniometer		Correlation[b]	Sig.[c]
	Median (IQR)[a]	Range	Median (IQR)	Range	R_s	(2-tailed)
Maximum passive wrist extension	56.1° (28.0)	-15.7° - 89.0°	36.6° (36.9)	-38.7° - 80.0°	0.772*	0.003*

[a] IQR = Interquartile Range, [b] Correlation Rr = Spearman rank correlation coefficient, [c] Sig. (2-tailed) = Two-tailed significance

* Correlation is significant at the level of 0.01 level (two-tailed)
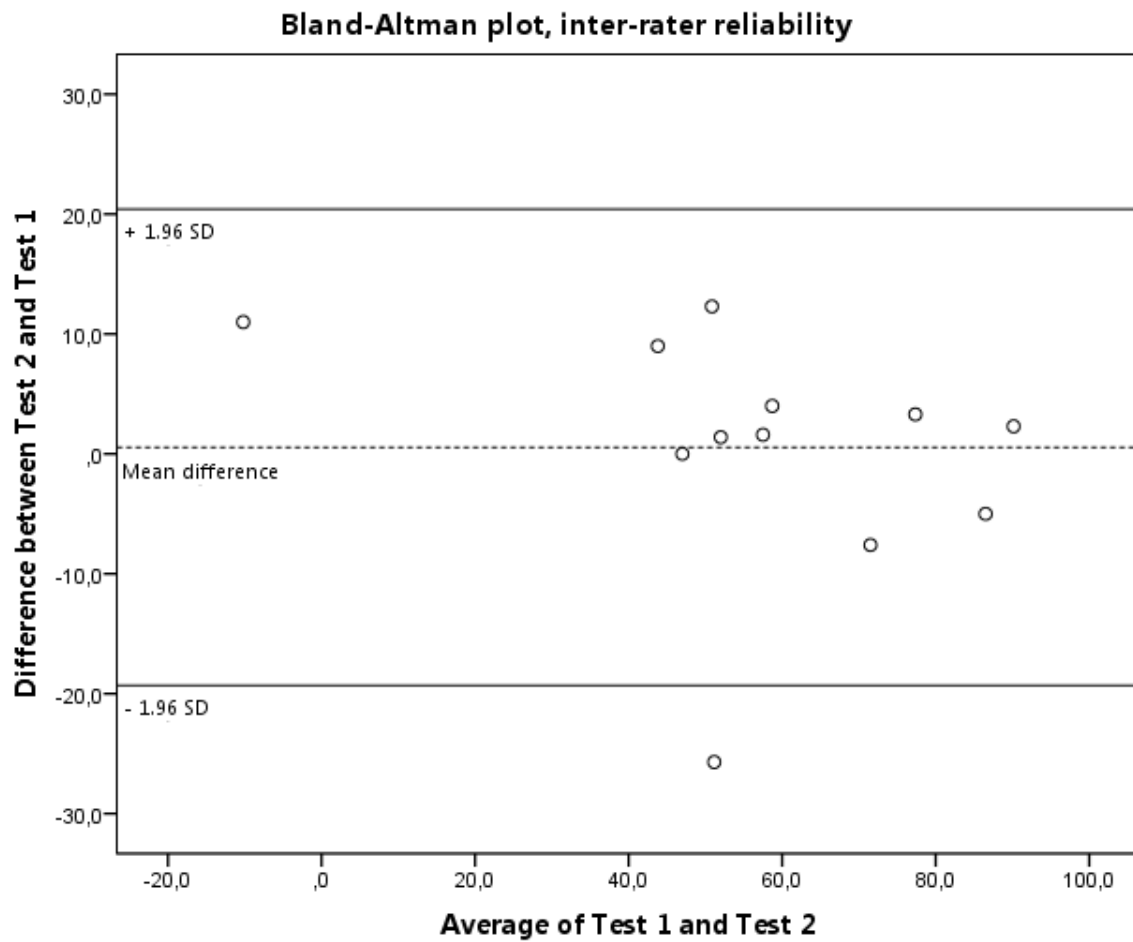
Fig. 1: Bland-Altman plot, inter-rater reliability based on the difference between Test 2 and Test 1 against the average of Test 1 and Test 2, measurements of Test 1 and Test 2 were both obtained by 4 different raters
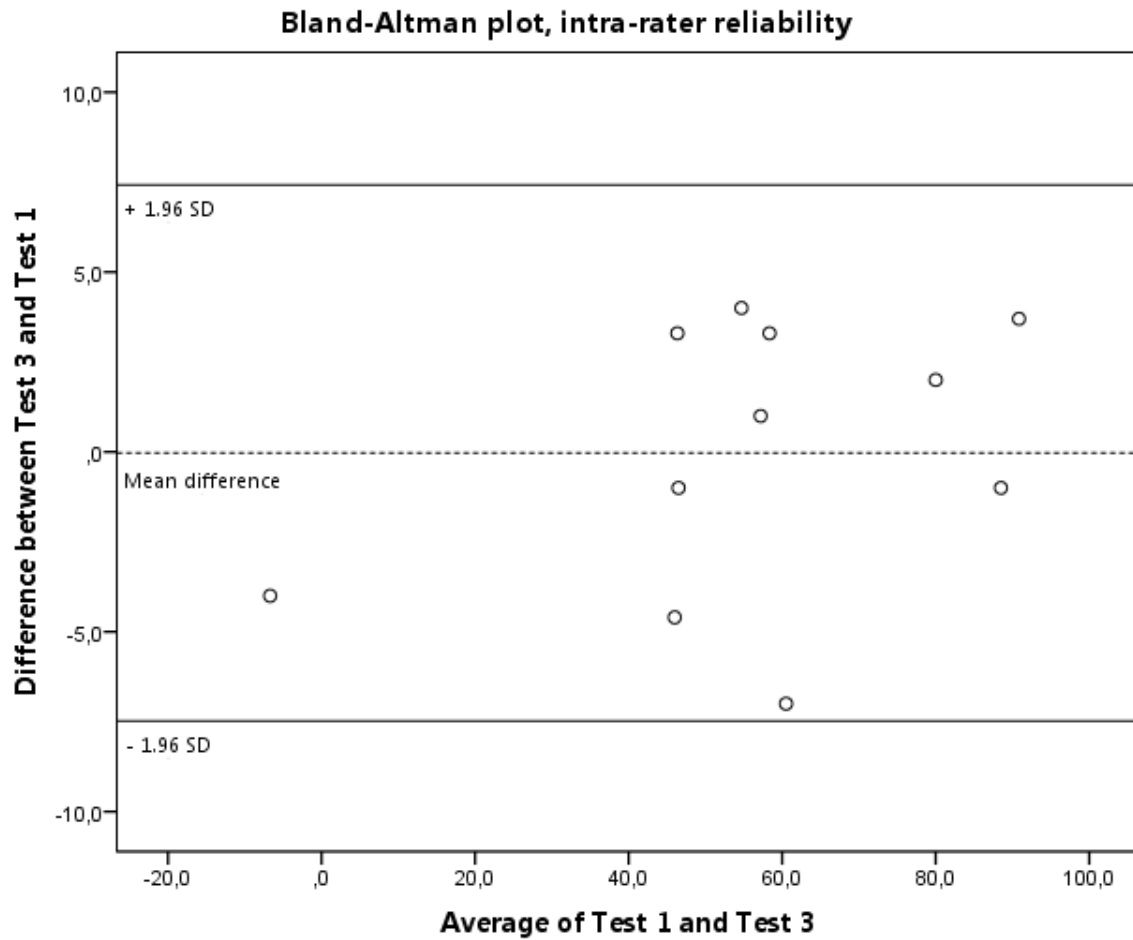
Fig. 2: Bland-Altman plot, intra-rater reliability based on the difference between Test 3 and Test 1 against the average of Test 1 and Test 3, measurements of Test 1 and Test 3 were both obtained by one, the same rater (NK)
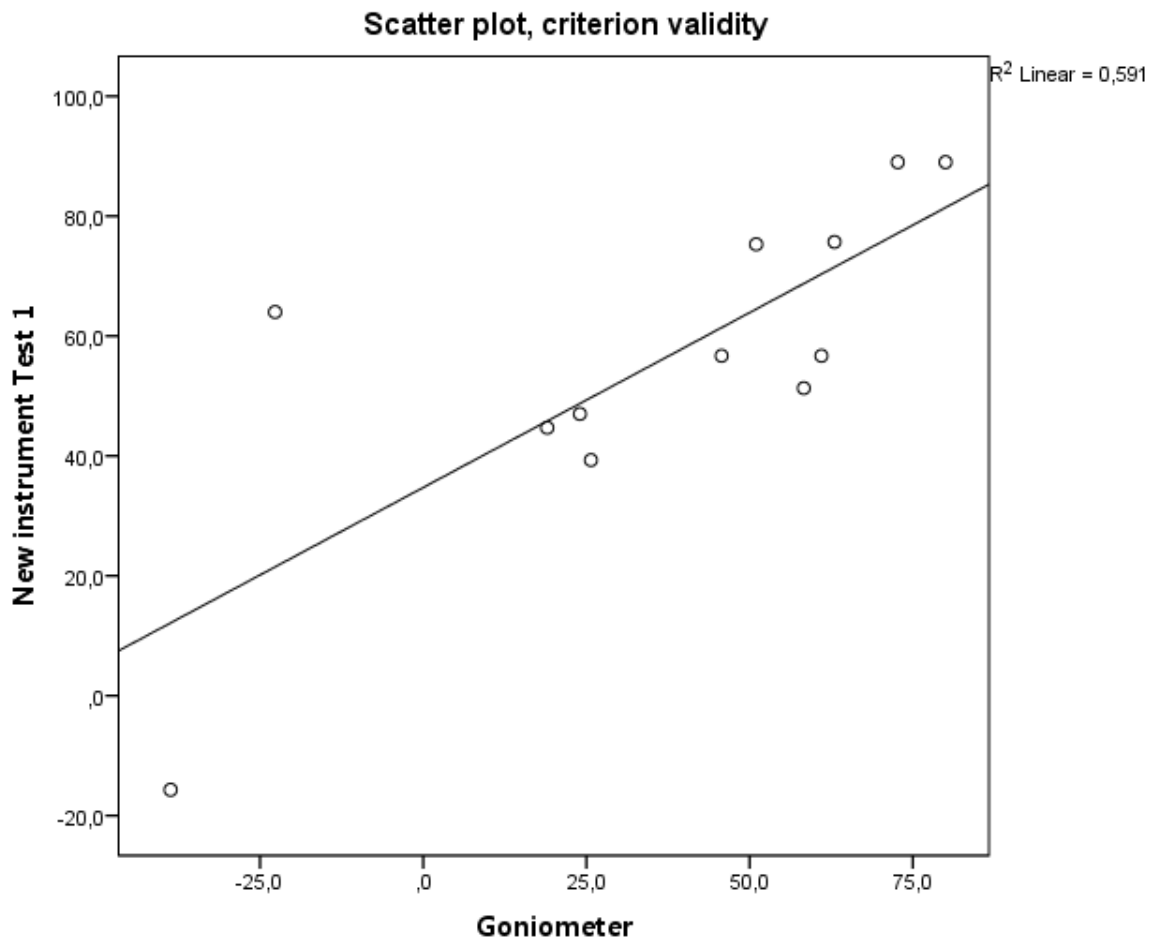
Fig. 3: Scatter plot, criterion validity based on the correlated results of the measurements obtained with the new instrument Test 1 and with the goniometer, measurements of Test 1 were obtained by four different raters and measurements of the goniometer by six raters

APPENDIX A

Fig.4: A new instrument to measure the length of the flexors of the forearm

APPENDIX B

Box 1: Description of different statistical analyzes

ICC [2.1] _{agreement}	$$\dfrac{BMS - EMS}{BMS + (k-1)\,EMS + \dfrac{k\,(RMS - EMS)}{n}}$$
ICC [3.1] _{consistency}	$$\dfrac{BMS - EMS}{BMS + (k-1)\,EMS}$$
BMS*	**B**etween-subjects **M**ean **S**quare
RMS*	Between **R**aters **M**ean **S**quare
EMS*	**E**rror **M**ean **S**quare (residual)
k	Number of raters (or ratings)
n	Number of subjects
LoA	Mean difference ± 1.96 x SD
$SEM_{agreement}$	$\sqrt{(\sigma^2_{pt} + \sigma^2_{residual})}$
$SEM_{consistency}$	$\sqrt{(\sigma^2_{residual})}$
σ^2_{pt}	Variance between raters = RMS
$\sigma^2_{residual}$	Error variance = EMS
SDC	$1.96 \times \sqrt{2} \times SEM_{agreement}$
SDC	$1.96 \times \sqrt{2} \times SEM_{consistency}$

* The values for BMS, RMS and EMS are taken from the ANOVA

Abbreviations: ICC: intraclass correlation coefficient, LoA: Limits of Agreement,

SD: Standard Deviation, SEM: Standard Error of Measurement, SDC: Smallest Detectable Change