Three-dimensional kinematics of shoulder function in stroke patients

Inter- and intra-rater reliability

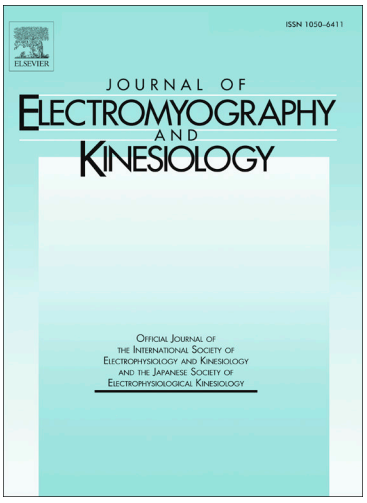
Hansen, Gunhild M; Kersting, Uwe G; Pedersen, Asger R; Svendsen, Susanne W; Nielsen, Jørgen F

Published in:
Journal of Electromyography & Kinesiology

DOI (link to publication from Publisher):
10.1016/j.jelekin.2019.05.006

Publication date:
2019

Document Version
Accepted author manuscript, peer reviewed version



Three-dimensional kinematics of shoulder function in stroke patients: inter- and intra-rater reliability

Gunhild M. Hansen, Uwe G. Kersting, Asger R. Pedersen, Susanne W. Svendsen, Jørgen F. Nielsen

PII: S1050-6411(19)30004-5
DOI: https://doi.org/10.1016/j.jelekin.2019.05.006
Reference: JJEK 2305

To appear in: *Journal of Electromyography and Kinesiology*

Received Date: 6 January 2019
Revised Date: 4 May 2019
Accepted Date: 7 May 2019

Please cite this article as: G.M. Hansen, U.G. Kersting, A.R. Pedersen, S.W. Svendsen, J.F. Nielsen, Three-dimensional kinematics of shoulder function in stroke patients: inter- and intra-rater reliability, *Journal of Electromyography and Kinesiology* (2019), doi: https://doi.org/10.1016/j.jelekin.2019.05.006

Three-dimensional kinematics of shoulder function in stroke patients: inter- and intra-rater reliability

Author names and affiliations

Gunhild M. Hansen [a,*], Uwe G. Kersting [bc], Asger R. Pedersen [a], Susanne W. Svendsen [a], Jørgen F. Nielsen [a]

[a] Hammel Neurorehabilitation Centre and University Research Clinic, Aarhus University, Voldbyvej 15, 8450 Hammel, Denmark

[b] Sport Sciences, Department of Health Science and Technology, Aalborg University, Fredrik Bajers Vej 7, 9220 Aalborg, Denmark

[bc] Institute of Biomechanics and Orthopaedics, German Sport University Cologne, Cologne, Germany

*** Corresponding author**

E-mail address: guhans@rm (G. M. Hansen)

Full postal address: Voldbyvej 15, 8450 Hammel, Denmark

Word count

Abstract: 196

Main text: 4090

Figures/tables: 3/2

ABSTRACT

Background: It is important to monitor progress during rehabilitation of stroke patients. To that end, clinical function tests may be supported by three-dimensional kinematic measures. The aim of this study was to evaluate the inter- and intra-rater reliability of three-dimensional kinematic measures of shoulder movements in stroke patients with reduced shoulder function.

Methods: Seventeen patients were tested in three sessions by two trained raters. Three-dimensional motion capture was performed of the more affected upper extremity and the trunk. Measures of movements of the scapula and humerus related to the trunk, the trunk related to the laboratory, the forearm related to the humerus, and temporospatial measures were obtained during two reach tasks from the Wolf Motor Function Test, ReachLow and ReachHigh. Inter- and intra-rater reliability was quantified with intraclass correlation coefficients (ICC).

Findings: In general, range of movements of scapula, shoulder, trunk and elbow and movement time and reach length showed high inter-rater reliability (ICC_{∞} 0.84-0.98) and intra-rater reliability (ICC_{∞}0.75-1.00), A minimum of five trials per task were required to achieve reliable ICC estimates.

Interpretation: Selected three-dimensional kinematic measures can be used reliably to evaluate specific movements of the shoulder in stroke patients with reduced shoulder function.

Keywords: CIMT; constraint-induced movement therapy; scapula; upper extremity

Three-dimensional kinematics of shoulder function in stroke patients: inter- and intra-rater reliability

Author names and affiliations

Gunhild M. Hansen [a,*], Uwe G. Kersting [bc], Asger R. Pedersen [a], Susanne W. Svendsen [a], Jørgen F. Nielsen [a]

[a] Hammel Neurorehabilitation Centre and University Research Clinic, Aarhus University, Voldbyvej 15, 8450 Hammel, Denmark

[b] Sport Sciences, Department of Health Science and Technology, Aalborg University, Fredrik Bajers Vej 7, 9220 Aalborg, Denmark

[bc] Institute of Biomechanics and Orthopaedics, German Sport University Cologne, Cologne, Germany

*** Corresponding author**

E-mail address: guhans@rm (G. M. Hansen)

Full postal address: Voldbyvej 15, 8450 Hammel, Denmark

Word count

Abstract: 196

Main text: 4090

Figures/tables: 3/2

ABSTRACT

Background: It is important to monitor progress during rehabilitation of stroke patients. To that end, clinical function tests may be supported by three-dimensional kinematic measures. The aim of this study was to evaluate the inter- and intra-rater reliability of three-dimensional kinematic measures of shoulder movements in stroke patients with reduced shoulder function.

Methods: Seventeen patients were tested in three sessions by two trained raters. Three-dimensional motion capture was performed of the more affected upper extremity and the trunk. Measures of movements of the scapula and humerus related to the trunk, the trunk related to the laboratory, the forearm related to the humerus, and temporospatial measures were obtained during two reach tasks from the Wolf Motor Function Test, ReachLow and ReachHigh. Inter- and intra-rater reliability was quantified with intraclass correlation coefficients (ICC).

Findings: In general, range of movements of scapula, shoulder, trunk and elbow and movement time and reach length showed high inter-rater reliability (ICC_{∞} 0.84-0.98) and intra-rater reliability (ICC_{∞}0.75-1.00), A minimum of five trials per task were required to achieve reliable ICC estimates.

Interpretation: Selected three-dimensional kinematic measures can be used reliably to evaluate specific movements of the shoulder in stroke patients with reduced shoulder function.

Keywords: CIMT; constraint-induced movement therapy; scapula; upper extremity

1 Introduction

Shoulder mobility plays an important role in overall upper extremity (UE) function during activities of daily living (Rundquist et al., 2012) and clinical function tests (e.g. the Wolf Motor Function Test (WMFT) (Wolf et al., 2001)) are frequently used to assess shoulder function in stroke patients.

We recently used selected tasks from the WMFT to assess shoulder function after constraint-induced movement therapy (CIMT) and found that a substantial part of the patients reached a positive outcome (Hansen et al., 2018). However, it is challenging to use clinical function tests to identify and evaluate specific changes in the process of recovery (Alt Murphy et al., 2015) and in particular to differentiate between compensatory and non-compensatory movements (Levin et al., 2015). Hence, it has been proposed that three-dimensional (3D) kinematic measures should be considered essential to distinguish compensation from restitution and should be developed accordingly (Kwakkel, 2017).

Several studies have reported reliability estimates of 3D kinematic measures of scapula positions and movements in healthy populations and people with shoulder impingement syndrome (Roy et al., 2007), children with cerebral paresis (Jaspers, Feys, Bruyninckx, Harlaar, et al., 2011; Lempereur et al., 2012; Mahon et al., 2017) and stroke patients using static (Pain et al., 2018) or dynamic assessments in shoulder flexion and abduction (De Baets, Jaspers, et al., 2013; De Baets, Van Deun, et al., 2013).

Assessment of 3D measures of scapula movements in stroke patients during shoulder flexion has shown reliable measures within one measurement session, but further methodological optimisation is required for multiple test sessions (De Baets, Van Deun, et al., 2013). Inter- and intra-rater reliability estimates are lacking for functional, dynamic reach movements. For a reliability study to be relevant, test administration must represent the clinical field or research context where the instrument will be used (Karanicolas et al., 2009). To evaluate shoulder outcomes after CIMT, it is therefore necessary to gain more knowledge of inter- and intra-rater reliability of 3D kinematic measures. Tasks performed during the measurements should reflect different movement directions and variations in speed and distance (Rodgers et al., 2017) and the marker placement should follow the International Standard of Biomechanics (ISB) (De Baets, Jaspers, et al., 2013; Wu et al., 2005).

The aim of this study was to evaluate the inter- and intra-rater reliability of 3D kinematic measures of dynamic shoulder function. We hypothesised that 3D kinematic measures of dynamic shoulder function, using palpation-based marker placement on defined anatomical landmarks, can reliably quantify shoulder function during clinically established reach tasks in stroke patients with reduced shoulder function.

2 Methods

2.1 Study design

We developed the test protocol, standardised the test procedures and determined the study size based on a pilot study with 3D measurements of shoulder movements in eight CIMT patients in 2013. The current reliability study was based on CIMT patients in the outpatient department at Hammel Neurorehabilitation Centre, Denmark, in the period 1 December 2015 to 31 August 2016. Patients included in CIMT were screened for eligibility to participate. To assess inter-rater reliability, the included patients were tested independently and in a randomised order by two raters (Rater 1 and Rater 2) on the first day. The raters were mutually blinded to each other's procedures except for tape markings to ensure equal set-up across sessions (see section 2.6). A one-hour break was included after the first session to secure no remaining marks on the skin before the second session. To assess intra-rater reliability, all patients were tested again by Rater 1 after a median of 3 days (range 1-5) and before the start of CIMT. Measurements were performed for both UEs, starting with the less affected side. Guidelines for Reporting Reliability and Agreement Studies were applied (Kottner et al., 2011).

2.2 Study patients

In accordance with the description of CIMT (Wolf et al., 2006), inclusion criteria for CIMT were 1) stroke more than three months ago, 2) age ≥18 years and 3) reduced function in the more affected UE following stroke, but ability to actively extend the more affected wrist, two fingers and thumb by more than 10˚, three times within one minute. Exclusion criteria were: 1) inability to cooperate in CIMT and 2) inability to perform an independent transfer to/from toilet. The following additional exclusion criteria were applied for the present study: 3) stroke >24 months ago, 4) normal clinically assessed shoulder function defined as full active range of movement (RoM) in shoulder flexion and normal isometric strength in abduction and external rotation when compared to the contralateral side and 5) inability to place the more affected hand on a 20.3 cm high box on a table while sitting (see the ReachHigh task described in section 2.6). Written informed consent was obtained from all participants. The study was approved by the Danish Data Protection Agency (number: 2012-58-006), while the Committee on Health Research Ethics of Central Denmark Region waived approval (number: 1-10-72-215-15).

2.3 Raters

Two trained physiotherapists performed the 3D measurements. Rater 1 was a clinical specialist in neurological physiotherapy with approximately 15 years of experience in neurorehabilitation. Rater 2 had >10 years of experience of movement analysis in stroke patients.

2.4 3D test instruments

3D data of UE and trunk movements were collected with an eight-camera motion capture system (Qualisys Oqus 400+, Qualisys, Göteborg, Sweden) and tracked using Track Manager (Version 2.15, Qualisys, Göteborg, Sweden). Tests were performed in a movement laboratory using a table and chair designed and adjusted to match the clinical setup of the WMFT.

2.5 Test procedures

Before execution of the test protocol, clinical examination was performed to obtain a Fugl-Meyer proximal score, which assesses UE function in 18 tasks on a 3-point scale (0-2), yielding a summation score ranging from 0 (worst) to 36 (best) (Lee et al., 2015). Additionally, shoulder pain intensity at rest and during activity was registered using an 11-point numerical rating scale (NRS-11) where the verbal anchors were 0=no pain and 10=worst imaginable pain (Dunning et al., 2015). Pain was defined as an NRS-11 ≥1. A palpation guideline was used to ensure accurate location of anatomical landmarks (Sint Jan, 2007). Calibration of the motion capture system was performed a maximum of four hours prior to testing.

2.6 Test protocol

Thirty-three markers (12 mm) were placed on the scapula and arm segments, head, trunk and pelvis in a neutral standing position: arms hanging down and thumbs pointing anteriorly in the sagittal plane (German neutral position) (**Fig. 1a**). Markers were placed on the scapula before the other segments. To ensure correct placement of elbow markers, the patients laid their hands with palms down on a firm box adjusted in height to achieve approximately 90˚ flexion of the elbow joints. Tests began with two static capture periods (standing), followed by measurements during ten repetitions (trials) of two tasks which were ReachLow (WMFT task 5) and ReachHigh (WMFT task 6). These functional tasks were chosen because they did not require grasp and manipulation (Hansen et al., 2018). Each patient was seated on a 45.7 cm high back-less chair facing a 73.5 cm high table. In ReachLow the patient moved his/her hand from the ipsilateral thigh to a proximal location on the table (**Fig. 1b**) and in ReachHigh from a proximal location on the table to a distal location on a box 20.3 cm above the table (**Fig. 1c**). The back legs of the chair were placed 36.0 cm from the front edge of the table. Patients were instructed to start with the ReachLow task. Before the ReachHigh task, the chair-position could be adjusted to make it possible for the patient to reach

the goal with the less affected UE without the trunk leaning forward. Such deviations from the standardised position were marked with tape to ensure equal set-up across sessions. Instruction and demonstration of task execution were given before the first trial. Patients were instructed to perform each trial as fast as possible. Each trial was initiated by the rater who started the motion capture and gave the verbal instruction: “ready, set, go”. The rater recognised capture stop when the patient's hand was placed on the end destination. Ten trials for both UEs and for each of the two tasks were performed, producing altogether 2 x 10 trials per UE per session per patient. Only trials executed by the more affected UE were included in the present study. There was a planned rest period of 20-30 seconds between each trial, but the rest period could be prolonged due to technical issues and need for additional instruction. Any deviations from the test protocol were documented for each session. The marker model for the scapula included three markers on the locations recommended by Wu et al (Wu et al., 2005). It was chosen to not use a marker triad as proposed in other studies, e.g., Grewal et al (Grewal et al., 2017) to keep preparation time short and reliable. The marker placement was further based on the recommendations from Bourne et al (Bourne et al., 2011) which provides comparable accuracy and reliability than bone pin based kinematics for forward reaching. Their observations were considered in the tracking algorithm which took markers on the spine of the scapula as main tracking markers with the marker on the angulus inferior to contribute to the anterior tilt only.

2.7 Range of movement and temporospatial variables

Ten outcome variables were chosen based on the literature (De Baets, Van Deun, et al., 2013, 2016) and reliability estimates from pilot data analysis. Primary outcomes were RoM of scapula upward rotation, inward rotation and tilt (**Fig. 1d**), RoM of shoulder flexion and temporospatial measures, i.e., movement time and reach length (=length of wrist movement). Secondary outcomes were RoM

of trunk lateral flexion, forward flexion and rotation and elbow extension. RoM was defined as the absolute difference between angular positions at maximum and minimum during each trial.

2.8 Data management and kinematic evaluation

Rater 1 managed the kinematic data. A custom-built shoulder kinematic model was assigned in Visual3D (Wu et al., 2005), where RoMs were assessed by marker-based video recordings and further processed (C-Motion. Visual3D Professional v6.01.18, Germantown, USA). Seven segments were included in the kinematic model: scapula, humerus, forearm, hand, head, trunk and pelvis. Trunk angles were calculated relative to the laboratory coordinate frame. Coordinate systems followed ISB recommendations (Wu et al., 2005). First, the xiphoid process was included in a static model to establish a trunk joint coordinate system, but due to challenges of marker occlusion during dynamic trials, the xiphoid process was excluded from the dynamic model (De Baets, Jaspers, et al., 2016). Movement start was determined visually as the time when the wrist marker left the start position (i.e., the thigh or the proximal low location on the table) and movement stop was defined visually as the landing of the wrist marker on the end destination (proximal low location on the table or distal high location on the box in front) (**Figs. 1b and 1c**). Joint rotation orders were Y-X-Z for scapula relative to the trunk and Y-X-Y for humerus relative to the trunk (Wu et al., 2005). The participants did not practice task execution before recording leading to a concern of incomparable test-conditions across sessions. Therefore, we excluded the first and second trials of each task in each session from the analyses, leaving 8 trials for the more affected UE per task per session per patient. We also excluded trials with occlusions of wrist or scapula markers and trials where the hand touched the table during task execution. Marker data were filtered at 3-5 Hz prior to calculations (Levin et al., 2015). All data were time-normalised (De Baets, Van Deun, et al., 2013).

2.9 Statistics

Sample size calculation was based upon data from the pilot study and estimated by intraclass correlation coefficients ($ICC_{agreement}$) calculated from a two-way random effects ANOVA. With a true $ICC_{agreement}$ assumed to be 0.75, the study needed 17 patients to determine the $ICC_{agreement}$ with a 0.30 width of the confidence interval (CI) and a 5% significance level (stata sampicc).

Inter- and intra-rater reliability of the chosen variables was quantified by the ICCs between raters (inter-rater; Rater 1 versus Rater 2 on the same day) and within rater (intra-rater; Rater 1 on two days). The ICCs were estimated using a linear mixed model (SAS Institute Inc, n.d.) (procedure MIXED) that separates variation in measurements into variation between patients (ω^2), variation between raters (inter-rater) or days (intra-rater) generally (θ^2) and within patients (τ^2) and variation between trials within patient and rater (inter-rater) or days (intra-rater) (σ^2); σ^2 will also be termed trial variance. Linear mixed models require normally distributed data. To fulfill this, the kinematic data were transformed by the square root function and the temporospatial measures by the natural logarithmic function before analysis.

Trial variance heterogeneity was pronounced (**Fig. 2**). Therefore, patterns of trial variance heterogeneity were analysed and tested by the generalised linear model with log-link and gamma distribution of the trial variances (McCullagh and Nelder, 1989; SAS Institute Inc, n.d.) (procedure GENMOD), and identified patterns were then included in the linear mixed model with trial variances fixed at their respective sample variances (pooled in case of homogeneity between raters, days or patients). Six of the 40 considered variables (ten variables for ReachLow, ten for ReachHigh and both inter- and intra-rater analyses) showed complete trial variance homogeneity, 25 showed trial variance heterogeneity between patients, but not between raters or days within rater,

whereas the remaining nine variables showed trial variance heterogeneity both between patients in general and between raters or days within patients.

By definition, the ICC is the correlation between measurements on the same patient made by different raters (inter-rater) or on different days by the same rater (intra-rater), where the measure is the average of trial values. In the ideal case of complete trial variance homogeneity, and the same number of trials, n, in all measurements, the ICC formula is:

$$ICC = \frac{\omega^2}{\sqrt{\omega^2 + \theta^2 + \tau^2 + \frac{\sigma^2}{n}}}$$

However, when trial variances and numbers differ between patients and raters (inter-rater) or days (intra-rater), the ICC formula is:

$$ICC_p = \frac{\omega^2}{\sqrt{\omega^2 + \theta^2 + \tau^2 + \frac{\sigma_{p1}^2}{n_{p1}}}\sqrt{\omega^2 + \theta^2 + \tau^2 + \frac{\sigma_{p2}^2}{n_{p2}}}}$$

Here subscripts 1 and 2 denote rater (inter-rater) and day (intra-rater), and subscript p emphasises that the formula produces patient specific ICC values. For large trial numbers, the ICC_p approaches the ICC value corresponding to infinitely many trials (or no trial variation):

$$ICC_\infty = \frac{\omega^2}{\sqrt{\omega^2 + \theta^2 + \tau^2}}$$

The ICC_∞ was estimated with 95% confidence interval (CI) from the linear mixed model. Valid data from the three sessions was used to calculate mean estimates of medians for each variable. Visual inspection of estimated inter-rater ICC values for every variable according to the hypothesised number of trials was used to assess the smallest number of trials necessary for ICC_p to reach $ICC_\infty \geq 0.75$ when computed with the largest observed trial variance of each variable (variables displaying $ICC_\infty < 0.75$ were excluded from further analyses of the required number of

trials). For similar trial variances, this number could be used as a guideline for the required number of trials for obtaining ICC_p close to ICC_∞.

ICC_∞ estimates of inter- and intra-rater reliability were interpreted using the following scale: poor (<0.40), moderate (0.40-0.74) and high (≥0.75) (Salter et al., 2005).

3 Results

Sixty stroke patients were screened for CIMT eligibility. Nineteen patients could not be included in CIMT and ten CIMT patients did not fulfil the eligibility criteria for the current study. Eleven eligible patients could not be included due to practical considerations, e.g. transport or coordination of raters' time schedules. Of the originally 20 included patients, 17 were able to adhere to the test protocol and therefore entered the analyses (Table 1). We excluded 120 out of 816 trials (17 patients x 2 tasks per patient x 3 sessions per task x 8 trials per session) due to technical errors. 118 trials were excluded due to occlusions of wrist or scapula markers and 2 trials due to the hand touching the table during task execution. For each patient, 2-8 trials per session were analysed for each task with a median of 8 trials per session for each task; 13% of the patients were represented with 2-6 trials and 87% with 7-8 trials per session for each task. Verbal instruction for one patient with aphasia was adjusted from “ready, set, go” to “start” in all sessions. Chair position was adjusted in distance to table for ReachHigh in nine patients with a median of 2 cm (range 0-5 cm).

Table 2 presents inter- and intra-rater reliabilities in terms of ICC_∞ for the RoM and temporospatial variables in each of the two tasks. Some variables showed relatively high variation between patients, raters and days, while others showed relatively low variation as illustrated by two variables (**Fig. 3**, upper row). Five trials were estimated to be sufficient to reduce the influence of trial-by-

trial heterogeneity and reach an ICC_p ≥0.75 for variables with ICC_∞ ≥0.75 in the inter-rater values (**Fig. 3**, lower row).

Overall, inter- and intra-rater reliability showed a tendency towards higher ICC_∞ estimates in the ReachLow task than in the ReachHigh task, apart from the reliabilities for shoulder flexion and the intra-rater reliability for scapula inward rotation. Inter-rater reliability was generally high (ICC_∞ 0.84-0.98), apart from moderate estimates for shoulder flexion during ReachLow (ICC_∞ 0.66) and for scapula inward rotation (ICC_∞ = 0.68) and tilt (ICC_∞ = 0.67) during ReachHigh. Intra-rater reliability displayed high ICC_∞ for most RoM variables (0.75-1.00), although the reliability for scapula inward rotation (ICC_∞ = 0.68) and shoulder flexion (ICC_∞ = 0.74) was moderate during ReachLow and for scapula upward rotation during ReachHigh (ICC_∞ = 0.67). Inter- and intra-rater reliability for the temporospatial variables was high in both ReachLow and ReachHigh (ICC_∞ = 0.82-0.96).

4 Discussion

The results demonstrated that in general, RoMs and temporospatial measures of the more affected UE could be measured reliably in post stroke CIMT patients with reduced shoulder function. In the ReachLow task, high inter- and intra-rater reliability was found for scapula movements, trunk movements and elbow extension, except that scapula inward rotation showed moderate intra-rater reliability. Shoulder flexion showed moderate reliabilities. In the ReachHigh task, high inter- and intra-rater reliability was found for shoulder flexion, trunk movements and elbow extension. Scapula upward rotation showed high inter-rater reliability while scapula inward rotation and tilt showed high intra-rater reliability. The temporospatial measures showed high reliabilities in both tasks. Five trials were estimated to be the minimum number required to get reliable ICC estimates.

When assessing reliability in human movement studies, two sources of variation have been pointed out (Schwartz et al., 2004). Intrinsic variations may result from sources that contribute with natural variation between and within patients while extrinsic variations are typically caused by experimental uncertainties (Schwartz et al., 2004). Extrinsic variations could therefore be a result of limitations in the test protocol in the current study. Movement start and stop were assessed by Rater 1 during data management. Other studies have defined movement start as the time point at which the velocity of the hand exceeded for example 5% peak velocity and movement stop as the time point at which the velocity of the hand fell below 5% peak velocity (Roby-Brami et al., 2016; Sethi et al., 2013), which might have minimised variability between measures. Although there may have been trials, where shoulder movement was initiated before the wrist left the start position or continued after the wrist reached the stop position, our definition of movement time matched the clinical implementation of the WMFT and was therefore considered relevant for the measurements of dynamic shoulder function in this study. We acknowledge the rather wide estimates of 95% CI in some of the variables, e.g. scapular tilt, which implies that the results should be interpreted with caution.

Errors in marker placement were possible sources of variation. Markers were placed with the patient in a standing position as suggested by the ISB (Wu et al., 2005). Although the American anatomical position (thumbs pointing outward) is recommended by the ISB (Wu et al., 2005), the German anatomically neutral position was applied since this was considered less challenging for the stroke patients in our study. Some patients had difficulties in assuming a symmetrical posture, which may have contributed to imprecise marker placement. More precise marker placement might have been obtained, for example by means of magnetic resonance imaging (Konda et al., 2018) or ultrasonography (Lee et al., 2013). Such solutions were not feasible in the current study and are most likely unrealistic in most clinical settings because they are time consuming and costly.

In the current study, the patients were instructed to perform each trial as fast as possible as described in the WMFT (Taub et al., 2011). This may have led to smoother movements and less compensatory trunk displacement compared to self-selected pace (Massie and Malcolm, 2012), but the opposite may also be the case. E.g., faster reaching in stroke patients with spasticity may encourage the use of compensations (Mandon et al., 2016).

A technical aspect affecting data analysis was the filtering frequencies applied to the raw data. Previous studies used filtering cut-off frequencies ranging from 2-6 Hz (Johansson et al., 2017; Seth et al., 2016), while others did not report any filtering (Hsieh et al., 2016; Merdler et al., 2013). We observed some trials where individual markers showed higher residuals, which could be avoided when going down to 3 Hz, but in most trials we used a 5 Hz cut-off frequency.

On the second test-day, no measure was included to elucidate potential changes in shoulder function and pain that could have contributed to variation in results, but the test days were not far apart and CIMT was not initiated between the two test days, which limited this potential source of inflated inter-rater variability. Applying one more marker on the forearm could have minimised exclusion of trials due to occlusion of markers.

Scapula RoM variables did not reach high reliability in all movements. Active shoulder elevation ≥60° has previously been suggested as an absolute prerequisite to measure scapular movements (De Baets, Van Deun, et al., 2013). In the current study, the ReachLow task did not require shoulder elevation ≥60°, which may explain why we only found moderate intra-rater reliability for scapula inward rotation. However, we achieved high reliabilities for the remaining scapula measures in this task, which speaks against an absolute requirement of arm elevation ≥60°. For the ReachHigh, moderate reliability was found for intra-rater upward rotation and inter-rater inward rotation and tilt. Fatigue may have contributed to this finding; starting with ten trials of ReachLow could have

caused fatigue in stroke patients with reduced shoulder function, thus showing more variation in the ReachHigh task. A recent study assessing intra-rater reliability in stroke patients reported fatigue as a factor that could have contributed to decreased measurement reliability for scapula upward rotation (Pain et al., 2018). Variations in skin motion artefacts (because of scapula sliding under the skin during movements of the shoulder) may also have played a part, which could have been more prominent in the ReachHigh task (Roy et al., 2010).

Shoulder flexion showed higher ICC estimates in the ReachHigh task than in the ReachLow task, although the ReachLow task could be considered easier to perform for stroke patients due to lower demands for dynamic control of the shoulder, compared to the ReachHigh task. However, the ReachLow task required the ability to reach forward without the hand touching the table, which was challenging for some patients. Adjustments of the hand placement in the start position could also have contributed to only moderate ICC estimates for shoulder flexion during ReachLow. A more exact description of the start position of the hand on the ipsilateral thigh could have contributed to more uniform task executions and thereby less variation (Jaspers, Feys, Bruyninckx, Cutti, et al., 2011). We chose not to strap the patients onto the chair because we wanted to measure compensatory movements of the trunk rather than restrict trunk movements during execution of the task.

The reliability of 3D measures of dynamic shoulder function was tested in post stroke CIMT patients with reduced shoulder function. Our findings can therefore only be generalised to patients sharing these characteristics. In our opinion, the results from our study are encouraging and selected variables with moderate or high inter- and intra-rater reliability could be applied as outcome measures in clinical studies of interventions to improve shoulder function.

5 Conclusion

It is possible to obtain reliable 3D measures of dynamic shoulder function in post stroke CIMT patients with reduced shoulder function across raters and days, when using a standardised test protocol with palpation-based marker placement on defined anatomical landmarks.

Acknowledgement

We thank William Sloth and Jakob Hansen for invaluable clinical contributions during the preparation and conduction of the study.

Funding

This study was not supported by any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

References

Alt Murphy M, Resteghini C, Feys P, et al. (2015) An overview of systematic reviews on upper extremity outcome measures after stroke. *BMC Neurology* 15(1): 29. DOI: 10.1186/s12883-015-0292-6.

Bourne DA, Choo AM, Regan WD, et al. (2011) The placement of skin surface markers for non-invasive measurement of scapular kinematics affects accuracy and reliability. *Annals of Biomedical Engineering* 39(2): 777–785. DOI: 10.1007/s10439-010-0185-1.

De Baets L, Jaspers E, Desloovere K, et al. (2013) A systematic review of 3D scapular kinematics and muscle activity during elevation in stroke subjects and controls. *Journal of electromyography and kinesiology* 23(1). Elsevier Ltd: 3–13. DOI: 10.1016/j.jelekin.2012.06.007.

De Baets L, Van Deun S, Desloovere K, et al. (2013) Dynamic scapular movement analysis: is it feasible and reliable in stroke patients during arm elevation? *PLoS ONE* 8(11). United States: e79046. DOI: 10.1371/journal.pone.0079046.

De Baets L, Jaspers E and Van Deun S (2016) Scapulohumeral control after stroke: a preliminary study of the test-retest reliability and discriminative validity of a clinical scapular protocol (ClinScaP). *NeuroRehabilitation* 38(4). Netherlands: 359–370. DOI: 10.3233/NRE-161327.

De Baets L, Van Deun S, Monari D, et al. (2016) Three-dimensional kinematics of the scapula and trunk, and associated scapular muscle timing in individuals with stroke. *Human Movement Science* 48. Netherlands: Elsevier B.V.: 82–90. DOI: 10.1016/j.humov.2016.04.009.

Dunning J, Mourad F, Giovannico G, et al. (2015) Changes in shoulder pain and disability after thrust manipulation in subjects presenting with second and third rib syndrome. *Journal of*

Manipulative and Physiological Therapeutics 38(6). National University of Health Sciences: 382–394. DOI: 10.1016/j.jmpt.2015.06.008.

Grewal TJ, Cudlip AC and Dickerson CR (2017) Comparing non-invasive scapular tracking methods across elevation angles, planes of elevation and humeral axial rotations. *Journal of Electromyography and Kinesiology*. DOI: 10.1016/j.jelekin.2017.10.001.

Hansen GM, Svendsen SW, Brunner I, et al. (2018) Predicting shoulder function after constraint-induced movement therapy: a retrospective cohort study. *Topics in Stroke Rehabilitation* 9357. Taylor & Francis: 1–7. DOI: 10.1080/10749357.2018.1440508.

Hsieh Y, Liing R, Lin K, et al. (2016) Sequencing bilateral robot-assisted arm therapy and constraint-induced therapy improves reach to press and trunk kinematics in patients with stroke. *Journal of neuroengineering and rehabilitation* 13. England: 31. DOI: 10.1186/s12984-016-0138-5.

Jaspers E, Feys H, Bruyninckx H, Cutti A, et al. (2011) The reliability of upper limb kinematics in children with hemiplegic cerebral palsy. *Gait and Posture* 33(4): 568–575. DOI: 10.1016/j.gaitpost.2011.01.011.

Jaspers E, Feys H, Bruyninckx H, Harlaar J, et al. (2011) Upper limb kinematics: Development and reliability of a clinical protocol for children. *Gait and Posture* 33(2). Elsevier B.V.: 279–285. DOI: 10.1016/j.gaitpost.2010.11.021.

Johansson GM, Grip H, Levin MF, et al. (2017) The added value of kinematic evaluation of the timed finger-to-nose test in persons post-stroke. *Journal of NeuroEngineering and Rehabilitation* 14(1). Journal of NeuroEngineering and Rehabilitation: 11. DOI: 10.1186/s12984-017-0220-7.

Karanicolas PJ, Bhandari M, Kreder H, et al. (2009) Evaluating agreement: conducting a reliability study. *Journal of Bone and Joint Surgery - Series A* 91(SUPPL. 3): 99–106. DOI: 10.2106/JBJS.H.01624.

Konda S, Sahara W and Sugamoto K (2018) Directional bias of soft-tissue artifacts on the acromion during recording of 3D scapular kinematics. *Journal of Biomechanics* 73. Elsevier Ltd: 217–222. Available at: https://doi.org/10.1016/j.jbiomech.2018.03.005.

Kottner J, Audige L, Brorson S, et al. (2011) Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *International Journal of Nursing Studies* 48(6). Elsevier Ltd: 661–671. DOI: 10.1016/j.ijnurstu.2011.01.016.

Kwakkel G (2017) Standardised measurement of sensorimotor recovery in stroke trials: consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable (SRRR). *International Journal of Stroke* in press(5): 451–461. DOI: 10.1177/1747493017711813.

Lee S, Yang D, Kim H, et al. (2013) A comparison of 3D scapular kinematics between dominant and nondominant shoulders during multiplanar arm motion. *Indian Journal of Orthopaedics* 47(2): 135. DOI: 10.4103/0019-5413.108882.

Lee Y-Y, Hsieh Y-W, Wu C-Y, et al. (2015) Proximal Fugl-Meyer assessment scores predict clinically important upper limb improvement after 3 stroke rehabilitative interventions. *Archives of physical medicine and rehabilitation* 96(12). United States: 2137–2144. DOI: 10.1016/j.apmr.2015.07.019.

Lempereur M, Brochard S, Mao L, et al. (2012) Validity and reliability of shoulder kinematics in typically developing children and children with hemiplegic cerebral palsy. *Journal of*

Biomechanics 45(11): 2028–2034. DOI: 10.1016/j.jbiomech.2012.05.020.

Levin MF, Liebermann DG, Parmet Y, et al. (2015) Compensatory versus noncompensatory shoulder movements used for reaching in stroke. *Neurorehabilitation and Neural Repair* 30(7): 635–646. Available at: http://nnr.sagepub.com/cgi/doi/10.1177/1545968315613863.

Mahon J, Malone A, Kiernan D, et al. (2017) Reliability of 3D upper limb motion analysis in children with obstetric brachial plexus palsy. *Physiological Measurement* 38(3). IOP Publishing: 524–538. DOI: 10.1088/1361-6579/aa5c13.

Mandon L, Boudarham J, Robertson J, et al. (2016) Faster reaching in chronic spastic stroke patients comes at the expense of arm-trunk coordination. *Neurorehabilitation and neural repair* 30(3): 209–20. DOI: 10.1177/1545968315591704.

Massie CL and Malcolm MP (2012) Instructions emphasizing speed improves hemiparetic arm kinematics during reaching in stroke. *NeuroRehabilitation* 30(4): 341–350. DOI: 10.3233/NRE-2012-0765.

McCullagh PNJA and Nelder JA (1989) *Generalized linear models*. *CRC Press Vol. 37*. London: Chapman and Hall. DOI: 10.1007/978-1-4899-3242-6.

Merdler T, Liebermann DG, Levin MF, et al. (2013) Arm-plane representation of shoulder compensation during pointing movements in patients with stroke. *Journal of Electromyography and Kinesiology* 23(4). Elsevier Ltd: 938–947. DOI: 10.1016/j.jelekin.2013.03.006.

Pain LAM, Baker R, Sohail QZ, et al. (2018) Three-dimensional assessment of the asymptomatic and post-stroke shoulder: intra-rater test–retest reliability and within-subject repeatability of the palpation and digitization approach. *Disability and Rehabilitation* 0(0). Informa UK Ltd.:

1–9. DOI: 10.1080/09638288.2018.1451924.

Roby-Brami A, Robertson JVG, Roren A, et al. (2016) Kinematics of the shoulder girdle during pointing: coordination between joints and their contribution to the peri-personal workspace. *Motor Control*: 1–44. DOI: 10.1123/mc.2015-0076.

Rodgers H, Shaw L, Bosomworth H, et al. (2017) Robot assisted training for the upper limb after stroke (RATULS): study protocol for a randomised controlled trial. *Trials* 18(1). Trials: 340. DOI: 10.1186/s13063-017-2083-4.

Roy JS, Moffet H, Hebert LJ, et al. (2007) The reliability of three-dimensional scapular attitudes in healthy people and people with shoulder impingement syndrome. *BMC Musculoskelet Disord* 8: 49. DOI: 10.1186/1471-2474-8-49.

Roy JS, Moffet H, McFadyen BJ, et al. (2010) The kinematics of upper extremity reaching: a reliability study on people with and without shoulder impingement syndrome. *Sports medicine, arthroscopy, rehabilitation, therapy & technology : SMARTT* 2: 8. DOI: 10.1186/1758-2555-2-8.

Rundquist PJ, Dumit M, Hartley J, et al. (2012) Three-dimensional shoulder complex kinematics in individuals with upper extremity impairment from chronic stroke. *Disability and Rehabilitation* 34(5). England: 402–407. DOI: 10.3109/09638288.2011.607214.

Salter K, Jutai JW, Teasell R, et al. (2005) Issues for selection of outcome measures in stroke rehabilitation: ICF participation. *Disability and Rehabilitation* 27(9): 507–528. DOI: 10.1080/0963828040008552.

SAS Institute Inc (n.d.) SAS/STAT software, Version 9.4 for Windows. 9.4. Cary, NC, USA: COMPUTER software industry; SAS Institute Inc.; UNITED States.

Schwartz MH, Trost JP and Wervey RA (2004) Measurement and management of errors in quantitative gait data. *Gait and Posture* 20(2): 196–203. DOI: 10.1016/j.gaitpost.2003.09.011.

Seth A, Matias R, Veloso AP, et al. (2016) A biomechanical model of the scapulothoracic joint to accurately capture scapular kinematics during shoulder movements. *PLoS ONE* 11(1): 1–18. DOI: 10.1371/journal.pone.0141028.

Sethi A, Davis S, McGuirk T, et al. (2013) Effect of intense functional task training upon temporal structure of variability of upper extremity post stroke. *Journal of hand therapy : official journal of the American Society of Hand Therapists* 26(2). United States: Elsevier Inc.: 132–7; quiz 138. DOI: 10.1016/j.jht.2012.09.001.

Sint Jan S van (2007) *Color atlas of skeletal landmark definitions: guidelines for reproducible manual and virtual palpation*. Edinburgh: Churchill Livingstone.

Taub E, Morris DM, Crago J, et al. (2011) Wolf Motor Function Test (WMFT) manual. *UAB CI Therapy Research Group*: 1–31.

Wolf SL, Catlin PA, Ellis M, et al. (2001) Assessing Wolf Motor Function Test as outcome measure for research in patients after stroke. *Stroke* 32(7): 1635–1639.

Wolf SL, Winstein CJ, Miller JP, et al. (2006) Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke. 296(17): 2095–2104. DOI: 10.1001/jama.296.17.2095.

Wu G, van der Helm FCT, Veeger HEJ, et al. (2005) ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand. *Journal of Biomechanics* 38(5): 981–992. DOI: 10.1016/j.jbiomech.2004.05.042.

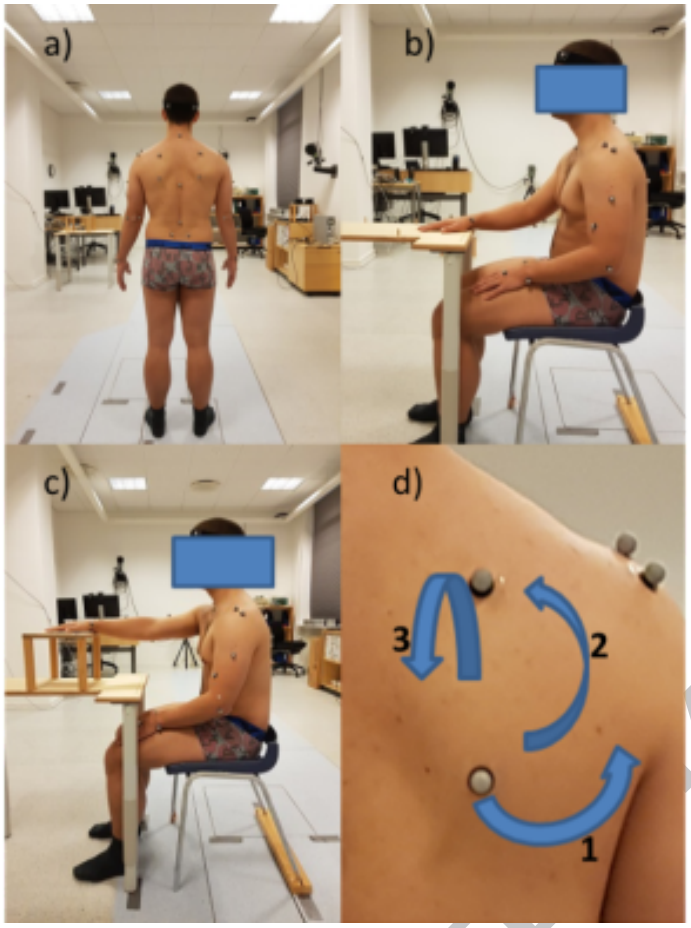
Fig. 1. Test setup and scapular movement directions; a) neutral standing position for marker placement, b) position for movement stop ReachLow and movement start ReachHigh, c) position for movement stop ReachHigh, d) right scapula movements (1=upward rotation, 2=inward rotation, 3=tilt).

Fig. 2. Variation in shoulder flexion measures during ReachLow. Patient-ID 1 Rater 1 (session 3) and Patient-ID 17 Rater 2 (session 1/2) show large variations. Patient-ID 3 Rater 1 (session 1/2) and Patient-ID 12 Rater 1 (session 3) show small variation.

Fig. 3. Inter-rater intraclass correlation coefficients for scapula upward rotation in ReachHigh and trunk rotation in ReachLow. The upper row illustrates variation between tasks, patients, raters and days. The lower row illustrates ICC_{∞} estimates by horisontal lines, while the grey areas show estimated ICC_p values according to the hypothesised number of trials and the observed range of trial variances in this study.

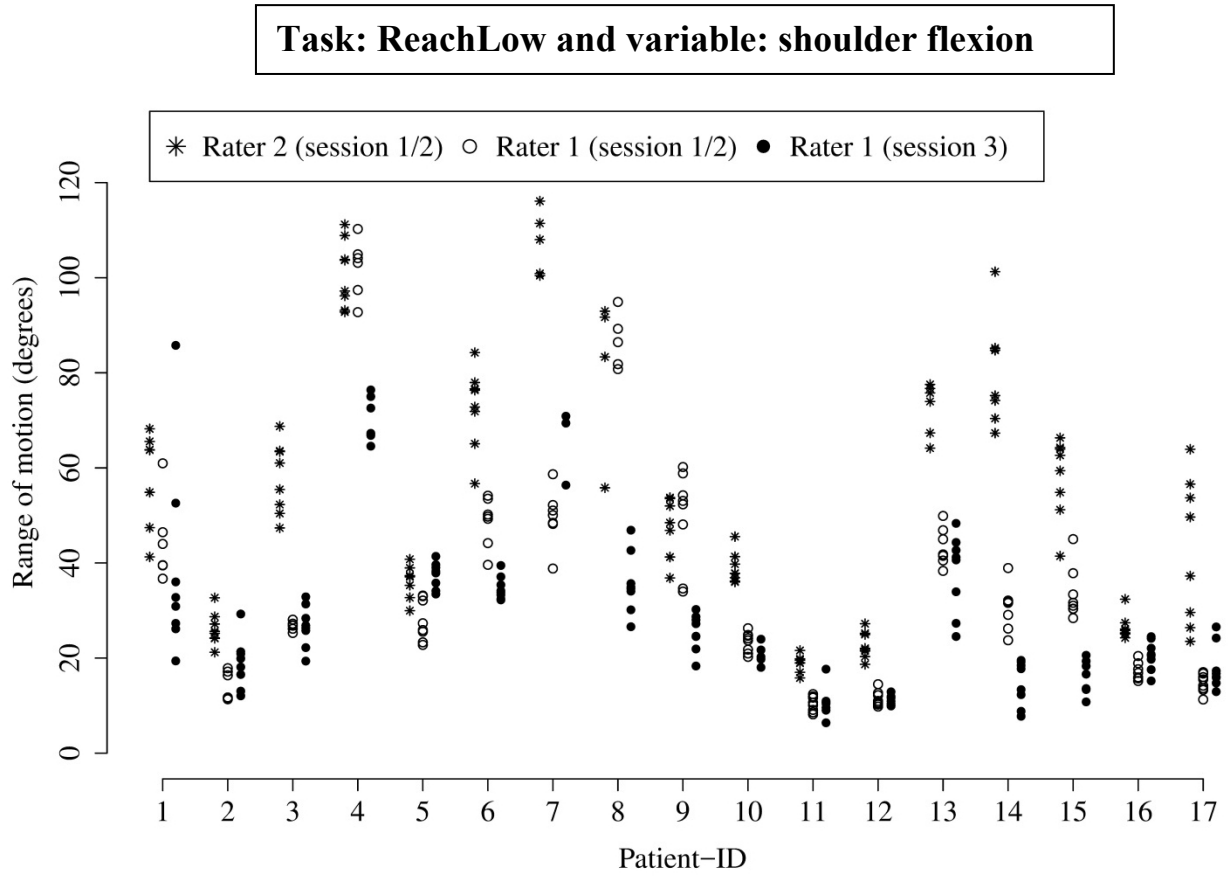
Gunhild M. Hansen

Fig. 1



Gunhild M. Hansen

Fig. 2



Gunhild M. Hansen

Fig. 3

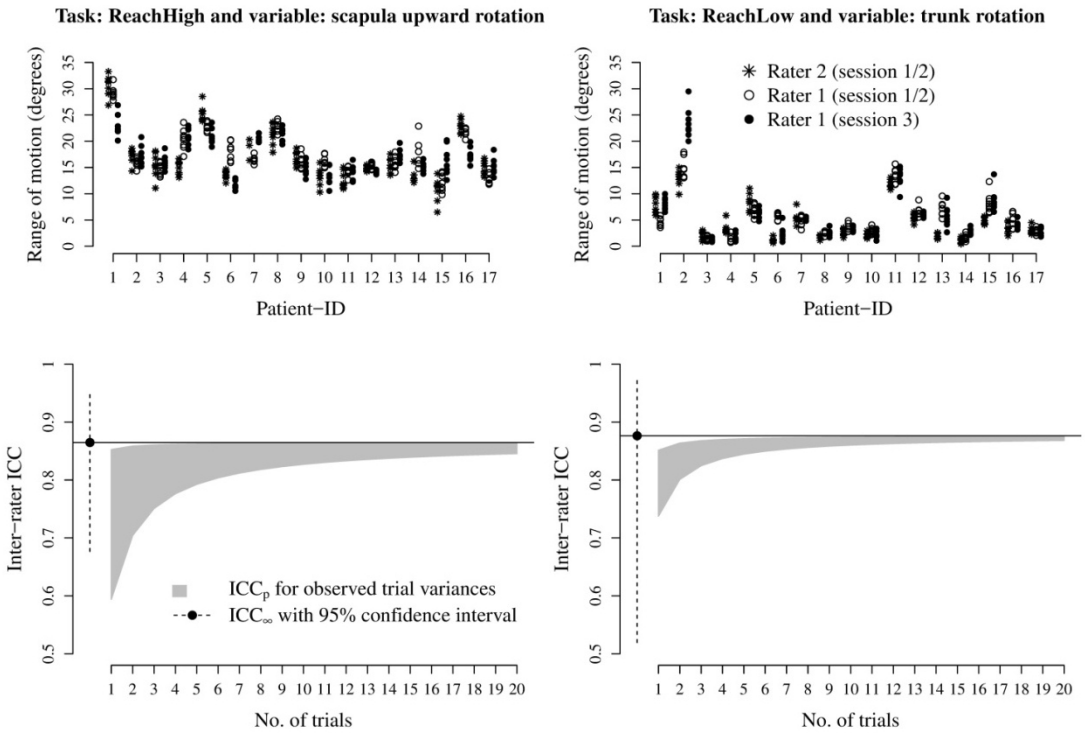


Table 1 Characteristics of the participating stroke patients, N=17.

Characteristics	**N**	**Mean**	**SD**
Sex, male/female	13/4		
Age, years		62.5	10.4
Time from stroke onset, days		181	72.8
Type of stroke, ischemia/haemorrhage	14/3		
First stroke, no/yes	3/14		
Localisation of stroke, supra-/infratentorial	13/4		
Dominant upper extremity, right/left	14/3		
More affected upper extremity, right/left	8/ 9		
Height, cm		175	7.8
Weight, kg		77.5	13.9
Fugl-Meyer proximal score *		20	5
Shoulder pain (NRS-11 ≥1) **			
At rest, no/yes	6/11		
During activity, no/yes	7/10		

* Fugl-Meyer proximal score: 0 (worst) - 36 (best).

** NRS-11: Numerical rating scale: 0 (no pain) - 10 (worst imaginable pain).

SD: standard deviation.

Table 2 Inter- and intra-rater reliability quantified with intraclass correlation coefficients, ICC infinity (ICC_{∞}), 95% confidence intervals (CI), and median values from the kinematic measures for primary and secondary outcomes in two tasks, ReachLow and ReachHigh.

		Rater 1 vs Rater 2 ReachLow			Rater 1 vs Rater 2 ReachHigh			Rater 1 vs Rater 1 ReachLow			Rater 1 vs Rater 1 ReachHigh		
		ICC_{∞}	95% CI	Median value	ICC_{∞}	95% CI	Median value	ICC_{∞}	95% CI	Median value	ICC_{∞}	95% CI	Median value
Primary outcomes	**Range of movement**												
	Scapula												
	upward rotation	0.89	0.73-0.96	8.8	0.86	0.68-0.95	17.1	0.76	0.49-0.91	9.0	0.67	0.36-0.88	17.2
	inward rotation	0.86	0.43-1.00	9.6	0.68	0.38-0.88	6.0	0.68	0.46-0.90	10.1	0.75	0.48-0.91	6.0
	tilt	0.87	0.31-0.99	4.2	0.67	0.37-0.88	3.2	0.86	0.65-0.95	4.4	0.75	0.39-0.94	3.3
	Shoulder flexion	0.66	0.20-0.94	17.8	0.92	0.80-0.97	42.4	0.74	0.46-0.90	18.4	0.92	0.80-0.97	30.2

	Temporospatial												
	Movement time	0.88	0.72-0.96	1.1	0.87	0.69-0.95	1.1	0.96	0.87-0.99	1.1	0.85	0.65-0.95	1.1
	Reach length	0.90	0.69-0.97	0.6	0.92	0.73-0.98	**0.0**0.5	0.91	0.75-0.97	**0.0**0.5	0.82	0.56-0.94	0.5
Secondary outcomes	**Range of movement**												
	Trunk												
	lateral flexion	0.93	0.81-0.98	5.3	0.98	0.05-1.00	6.3	0.86	0.66-0.95	5.2	0.81	0.57-0.93	6.2
	forward flexion	0.90	0.72-0.97	8.5	0.84	0.45-0.97	6.4	0.96	0.89-0.99	9.1	0.87	0.68-0.95	7.0
	rotation	0.88	0.52-0.98	4.5	0.92	0.81-0.97	5.0	0.91	0.78-0.97	5.0	1.00[1]		5.4
	Elbow extension	0.97	0.86-0.99	27.7	0.97	0.92-0.99	32.8	0.90	0.74-0.97	27.6	0.89	0.73-0.96	32.5

[1]The ICC_{∞} is estimated as 1.00 which is considered a high correlation, but it could also imply a large uncertainty of the estimation.

[2]Range of movement unit: degrees; movement time unit: seconds; reach length unit: centimeters.

Author Biography
04 May 2019

Gunhild M. Hansen is a physiotherapist and holds a Ph.D from Århus University, Denmark, in 2018, on the subject ”Shoulder function in stroke patients” measured with 3D kinematics and clinical outcome measures

Uwe G. Kersting is professor within the field of biomechanics at the Department of Health Science and Technology, The Faculty of Medicine Sport Sciences, Ålborg University, Denmark

Susanne W. Svendsen, dr med, is professor in the field of occupational medicine at Hammel Neurorehabilitation Centre and Research Clinic, Århus University, Denmark

Jørgen F. Feldbæk, dr med, is professor in neurological rehabilitation and leader of the Research Department at Hammel Neurorehabilitation Centre and Research Clinical, Århus University, Denmark