Responsiveness of kinematic and clinical measures of upper-limb motor function after stroke

DOI:
10.1016/j.rehab.2020.02.005

Document Version
Accepted author manuscript

Citation for published version (APA):
Villepinte, C., Verma, A., Dimeglio, C., Boissezon, X. D., & Gasq, D. (2020). Responsiveness of kinematic and clinical measures of upper-limb motor function after stroke: a systematic review and meta-analysis. *Annals of Physical and Rehabilitation Medicine*. https://doi.org/10.1016/j.rehab.2020.02.005

Published in:
Annals of Physical and Rehabilitation Medicine

Responsiveness of kinematic and clinical measures of upper-limb motor function after stroke: a systematic review and meta-analysis

Claire Villepinte OT MSc Arpana Verma MBChB MPH FFPH PhD
Chloe Dimeglio PhD Xavier De Boissezon MD PhD David Gasq MD
PhD

PII: S1877-0657(20)30059-2
DOI: https://doi.org/doi:10.1016/j.rehab.2020.02.005
Reference: REHAB 1366

To appear in: *Annals of Physical and Rehabilitation Medicine*

Received Date: 29 August 2019
Accepted Date: 15 February 2020

Please cite this article as: Villepinte C, Verma A, Dimeglio C, Boissezon XD, Gasq D, Responsiveness of kinematic and clinical measures of upper-limb motor function after stroke: a systematic review and meta-analysis, *Annals of Physical and Rehabilitation Medicine* (2020), doi: https://doi.org/10.1016/j.rehab.2020.02.005

Responsiveness of kinematic and clinical measures of upper-limb motor function after stroke: a systematic review and meta-analysis

Abstract

Background. Kinematic analysis and clinical outcome measures with established responsiveness contribute to the quantified assessment of post-stroke upper-limb function, the selection of interventions and the differentiation of motor recovery patterns.

Objective. This systematic review and meta-analysis aimed to report trends in use and compare the responsiveness of kinematic and clinical measures in studies measuring the effectiveness of constraint-induced movement, trunk restraint and bilateral arm therapies for upper-limb function after stroke.

Methods. In this systematic review, randomised controlled trials implementing kinematic analysis and clinical outcome measures to evaluate the effects of therapies in post-stroke adults were eligible. We searched 8 electronic databases (MEDLINE, EMBASE, Web of Science, Scopus, CINAHL, CENTRAL, OTseeker and Pedro). Risk of bias was assessed according to the Cochrane Risk of Bias domains. A meta-analysis was conducted for repeated design measures of pre- and post-test data providing estimated standardised mean differences (SMDs).

Results. We included reports of 12 studies (191 participants) reporting kinematic smoothness, movement duration and efficiency, trunk and shoulder range of motion, control strategy and velocity variables in conjunction with assessment by Motor Activity Log, Fugl-Meyer Assessment and Wolf Motor Function Test. Responsiveness was higher (i.e., non-overlap of 95% confidence intervals [CIs]) for Motor Activity Log score (SMD for amount of use 1.0, 95% CI 0.75–1.25, $p<0.001$; SMD for quality of movement 0.96, 95% CI 0.72–1.20, $p<0.001$) than movement efficiency, trunk and shoulder range of motion, control strategy and peak velocity.

Conclusion. These results are consistent with current literature supporting the use of combined kinematic and clinical measures for comprehensive and accurate evaluation of post-stroke upper-limb function. Future research should include other design trials and rehabilitation types to confirm these findings, focusing on subgroup analysis of type of rehabilitation intervention and functional levels.

Keywords: hemiparesis; upper extremity; kinematics; 3D motion analysis; outcome measure; psychometrics

Introduction

Upper-limb (UL) impairment due to neurological loss post-stroke typically includes upper motor neuron syndrome, somatosensory deficit contralateral to the brain lesion and cognitive disorders [1–3]. Studies estimate that 50% to 85% of individuals with acute stroke and 50% with chronic stroke present UL impairment [4–6], which influences functional ability in task accomplishment and independence, thereby affecting quality of life post-stroke [1,7–9].

Studies report a large number of measures used to evaluate UL function. In total, 53 measures identified contribute to a large diversity in the selection of measures across trials post-stroke [6,7,10,11]. Clinical guidelines recommend the use of valid, reliable and responsive measures contributing to evidence-based practice [10,12]. Studies show strong trends in the use of several measures to evaluate various International Classification of Functioning, Disability and Health (ICF) levels for UL function post-stroke [6,10,11,13,14]. The use of more than one measure for several ICF levels is reported in 72% of trials [6]. Discrepancies in the selection of measures are according to psychometric properties and/or intervention types [6,14]. Knowledge of psychometric properties of measures will contribute to potential comparisons of treatment effects [10,15], which will contribute to trial quality and allow for comparison of practices of studies via meta-analyses contributing to evidence-based care [6,10,13].

The use of psychometrically sound outcome measures contributes to the identification of appropriate rehabilitation interventions [3,6,7,10,13,16,17]. Generally, the evaluation of treatment effects in stroke trials depend on the use of observational measures of timed standardised tasks [10,18,19]. Measures assessing ICF activity-level function demonstrate limitations for discriminating motor recovery from motor compensation that is an important point to consider in constraint-induced movement therapy (CIMT), trunk restraint therapy (TRT) and bilateral arm therapy (BAT) [8,18,9]. Selection of appropriate interventions should consider whether outcomes are recovery- or compensatory-based for effects [8–10]. Recovery is defined as the return of physiological motor patterns with impaired structure, whereas compensatory movement refers to the ability to accomplish a task through adaptation of motor patterns [8,9,20].

Discrete strategies of movement and motor impairment may not be detected with clinical measures [6,21], which indicates a need for more precise measures of compensatory strategies or subtle impairments [8,19,22,20]. Furthermore, compensatory movement strategies, such as lateral flexion and forward displacement of the trunk in reaching patterns are not explicitly measured by most activity-level motor scales, with the exception of the Wolf Motor Function Test (WMFT) [8,22]. The use of kinematic measures combined with clinical measures aligned with ICF levels improves the ability to distinguish motor recovery from compensatory strategies [8,10,13,9].

The use of kinematic measures in UL evaluation is gaining recognition for the measurement of UL movement in stroke trials, supported by expert-based recommendations [6,13,22,23]. Combined use of kinematic motion analysis outcomes (KINOs) with clinical outcome measures (CLIOs) contributes to the differentiation of motor recovery and compensatory movement patterns as well as prognostic accuracy [6,8,10,22,20,24]. KINOs provide a quantitative and objective measure of UL body function level according to the ICF [10]. Furthermore, KINOs are well correlated with other clinical measures in stroke trials, demonstrating high to moderate significant correlation with Fugl-Meyer Assessment for upper extremity (FMA-UE), Action Research Arm Test (ARAT) and WMFT across reaching tasks [18,22,25]. The strength of correlations depends on the specific actions that are being measured and the variable measured in the task [8,22].

KINOs may present higher responsiveness to change than CLIOs [22,26]. KINOs for assessing UL post-stroke were present in 21 of 41 trials from 2010 to 2014 [22]. Studies report an increasing use of KINOs with CLIOs such as the FMA-UE [6] and with activity dexterity measures providing measurement of impairment of trunk movements in relation to UL tasks [10,27]. KINO variables have demonstrated established validity and reliability in stroke populations [10,22,28]; however, responsiveness remains understudied, reported in only 3 trials for movement duration, smoothness, and trunk displacement variables [22,26,28]. In addition, kinematic analysis has been based on small sample sizes affecting the generalizability of findings [22,23]. KINO variables may be more precise than clinical measures, with some studies demonstrating advantages of peak velocity and movement duration over timed performances of WMFT. However, these variables measured different aspects of UL movement [10,18]. Other studies demonstrated the ability of KINOs to provide higher precision of quantitative measurement of specific movement aspects over global clinical scales assessing broad constructs of UL function [22,25].

Recent recommendations encourage standardisation of kinematic analysis measurement protocols according to ICF levels, tasks and conditions analysed as well as psychometric properties [10,15,23]. However, evidence regarding KINO responsiveness remains limited [22].

The aim of this systematic review was to establish and compare KINO and CLIO responsiveness for UL function after post-stroke interventions. Our review sought to first report the effect sizes of KINO and CLIO changes related to CIMT, TRT, and BAT rehabilitation interventions to conduct the comparison.

Methods

This protocol was registered on the PROSPERO database on March 9, 2016 (registration no. CRD42015023907).

Literature search

Studies published in English and French were searched in electronic bibliographic databases, subject-specific databases, and unpublished grey literature, and authors were contacted to identify additional studies. Also, reference lists were manually scanned by 2 reviewers independently. The databases searched were MEDLINE, EMBASE, Web of Science, Scopus, CINAHL, Central Register of Controlled Trials (CENTRAL), OTSeeker and PEDro. No date limit was imposed in accordance with recent developments of KINOs [22]. The last search was performed on May 1, 2016, and the literature search was updated from May 1, 2016 to October 28, 2019.

Specific strategies were used for each database. The MEDLINE strategy is illustrated in Appendix A, supplementary methods. Index terms used were stroke, upper limb, kinematic and outcome measure. Methodological filters were used for randomised controlled trial (RCT) designs contributing to a sensitivity and precision maximising-based search [29].

Selection criteria

Studies were selected according to predetermined eligibility criteria based on Population-Interventions-Comparator-Outcomes-Study design(s) (PICOS) components [29,30].

Type of participants: Studies of adults ≥ 18 years old with UL impairment post-stroke who underwent CIMT, TRT and/or BAT interventions were considered. We excluded data for participants who received other therapies such as functional electric stimulation, robot and/or virtual reality therapy to ensure the homogeneity of groups for meta-analysis [29]. Neurological conditions other than stroke were excluded. No exclusion regarding time post-stroke was applied [20].

Types of intervention: Only studies implementing KINOs of UL movement with 3D optoelectronics motion capture systems were considered because of the high measurement accuracy of free movements, congruent with recommendations in stroke rehabilitation [22,23,31]. Trials with KINOs from virtual, robotic or haptic devices were excluded because of the difficulty in comparing data and/or restriction of UL movements [22,31]. We selected studies reporting KINOs according to established classifications, including trunk range of motion (ROM), shoulder ROM, elbow ROM, movement duration, peak velocity, movement efficiency (index of curvature in distal trajectory expressed as a path ratio), smoothness of movement (number of movement units or peaks on the velocity curve in distal, the score inversely proportional to smoothness), and control strategy (time of peak velocity expressed as a percentage of the movement duration) [22,31]. Studies were excluded if they did not report at least one of these measurement variables.

Types of comparators: Only studies implementing KINOs in conjunction with CLIOs for UL function post-stroke were included. CLIOs selected were the most frequently used with high-level psychometric properties and established clinical utility: FMA-UE, WMFT, ARAT, Chedoke Activity Hand and Arm Inventory (CAHAI), Motor Activity Log (MAL) and Box and Block Test (BBT) [6,7,10,11]. Studies not reporting at least one of these measures were excluded.

Types of outcome measures: Responsiveness of KINOs and CLIOs (described above) was computed from pre- and post-test means and standard deviation values, as a standardised mean difference (SMD) appropriate for continuous outcomes [29,32].

Study design and setting: Randomised trials reporting at least 2 measurement points (i.e., pre/post) were included. Studies conducted within clinical, hospital, rehabilitative-based and/or community settings were included because they contribute to external validity and hence the generalisability of the results [29].

Study selection and data extraction

Selection of studies was performed independently by 2 reviewers (CV and DG) after screening titles and abstracts. Studies were included after agreement by reviewers, with discrepancies resolved by a third author (AV). Duplicate publications were removed after contact with authors and application of a decision tree [30]. The collection and management of studies was conducted with Zotero (https://www.zotero.org/) referencing storage and spreadsheet software (Microsoft Office Excel) [33]. The data extraction form was pre-piloted [29] and used independently by 2 reviewers (CV and DG). Information extracted in accordance with PICOS included characteristics of participants, type of intervention, modalities of rehabilitation (Table 1), technical aspects of the kinematic evaluation (Table 2), KINOs and CLIOs (Table 3).

Risk of bias

Risk of bias of studies was determined independently by 2 reviewers (CV and DG) according to the Cochrane Collaboration Risk of Bias tool [29,30]. Judgement of bias according to the extracted information was determined as high, low or unclear risk due to lack of information [29]. Data were extracted by using RevMan 5.3 software. Disagreement between reviewers was resolved by the arbitrary decision of a third author (AV).

Data analysis

Meta-analyses were performed for continuous summary measures from individual studies with the SMD calculated by Hedges' g with 95% confidence intervals (CIs), p values and z scores [29,34]. This was appropriate for different measurement units for comparison between studies [29,30]. Hedges' g was computed for all outcomes as proposed by Borenstein [34] for paired continuous outcomes, considering the variance of pre- and post-test data. The sign of the post–pre differences was inverted for outcomes in which improvement is represented as a decrease in score (WMFT Time subscale; movement smoothness, duration and efficiency; and trunk ROM) to obtain a positive effect size related to the interventions. Because of lack of availability of correlation (r) data between pre- and post-test for studies, the pooled effect sizes for r = 0.5 were selected as previously proposed [34]. We performed a sensitivity analysis using a range of plausible correlations (r = 0, 0.25, 0.5, 0.75 and 0.9) for imputation of r values for all outcomes, allowing for analysis of the effect of association strength between pre- and post-test data on the between-studies effect size and between-outcomes responsiveness [29,34,35].

Meta-analysis of Hedges' g was performed for each KINO and CLIO across the studies, independent of the intervention (i.e., CIMT and/or TRT or BAT). Effect sizes were classified as small (<0.2), moderate (0.2–0.8) or large (>0.8) [30,34]. Heterogeneity or between-study variation of studies was determined by the I^2 statistic (low if <50% [29,36]) and visual assessment of forest plots [29]. A random effects model was systematically applied given the heterogeneous nature of the data collection and the large variance of the I^2 values obtained [37,38]. Forest plots were examined for similarities of the direction of treatment effects [39]. Meta-analysis was performed using the package *Metafor* for R software [40,41].

Comparison of responsiveness between outcomes involved classification according to the decreasing value of the pooled Hedges' g values. With a conservative hypothesis, we considered that a difference in effect size between 2 outcomes was significant in cases of no overlap of their 95% CIs.

Results

Study selection

A total of 479 articles were identified by the selection procedure presented in Figure 1. After adjusting for duplicates, 347 studies remained, for which titles and abstracts were screened; 332 studies were discarded. The full text of 15 studies was assessed for eligibility: 2 studies did not meet the inclusion criteria in terms of participants [19,25]; one study was a synthesis of 3 RCTs lacking individual results [43]; and a final study was a duplicate study [25,44]. An additional article was included from the manual search of reference lists [45]; no relevant unpublished studies were retained [30]. Authors of identified studies were contacted; however, this did not generate any additional trials. Finally, a total of 12 studies from 2006 to 2015 were identified, some studies including 2 distinct groups in terms of rehabilitation (CIMT and/or TRT or BAT).

Characteristics of the participants

Characteristics of participants in included studies are shown in Table 1. The mean age of participants was 56.6 years and the mean time since stroke onset was predominately chronic (>6 months) [20] across all trials except one [46]: mean 25.7 months (range 11.9 weeks to 21.5 months). Most studies included participants with mild (57 to 66) to moderate (28 to 57) severity of UL impairment (mean FMA-UE 46.9/66) [47].

Included studies involved 247 participants who underwent UL rehabilitation, including 128 for CIMT in 8 studies [5,46,48–53], 30 for TRT in 3 studies [45,54,55], 38 for BAT in 2 studies [51,56], and 51 for CIMT combined with TRT in 4 studies [5,49,52,53].

Intervention characteristics

Technical aspects of kinematic evaluation are in Table 2. Movement was recorded at spontaneous velocity across all studies except for one measuring maximal velocity [53], and 2 studies recorded both (data were analysed at spontaneous velocity) [51,56]. UL evaluation was assessed across unimanual functional tasks including reaching, point-to-point or reach-to-grasp with objects (a bell or can) [5,45,46,48–56], and 3 trials analysed bimanual tasks in both reaching and activities of daily living (opening a box or drawer and retrieving a note; not retained for data analysis) [50,51,56]. The evaluations retained for data analysis were systematically performed at the end of the rehabilitation programme.

Kinematic outcomes

KINOs are shown in Table 3. Trunk ROM was assessed in 7 trials but reported in 6 describing TRT with or without additional CIMT [5,45,49,52–54] (data not usable from the study by Michaelsen et al., 2006). Values were calculated from sagittal trunk displacement and often standardised by hand movement indicating the ratio of trunk to UL movement. Data from one study was not included [53] because of calculation according to 3 phases of the reaching movement. Five studies measured shoulder ROM: 2 after TRT [54,57] and 3 CIMT+TRT studies [49,52,53] with a predominance of shoulder flexion. Seven studies assessed elbow ROM after TRT, CIMT and CIMT+TRT, including one for which the data were not usable [55]. Four reports described metrics from 2 subgroups of TRT [45,54], 1 subgroup of CIMT [46], and 1 subgroup of CIMT/BAT [51]. Movement efficiency, movement duration and control strategy were the most reported, in 7 trials [5,46,49,54–56]. Smoothness was measured in 6 studies [46,48,49,51,54,55] among all intervention subgroups.

Clinical outcomes

CLIOs are shown in Table 3. Three studies reported the use of only one CLIO [48,50,54], either FMA-UE or MAL. Use of more than one ICF level measure was reported in the remaining studies. FMA-UE impairment measure, for all intervention subgroups, was used in 7 studies, but the data were reported in only 6 trials, for which 5 reported the combined use with an activity level measure including MAL, WMFT and/or BBT. All 9 studies reporting CIMT included MAL activity measures [5,46,48–52,56], used with other impairment and activity level measures in 7 studies. The activity measure WMFT was used in 3 studies [5,49,51] across all interventions. ARAT and BBT were the least reported measures for CIMT+TRT and TRT, respectively. No studies reported the use of CAHAI.

Risk of bias within studies

Risk of bias for included studies is in Figure 2. A detailed description of risk of bias for individual studies is in Appendix A, supplementary results – risk of bias in studies.

Results of individual studies

Raw values and results of meta-analyses for KINO and CLIO summary measures from individual studies (SMD calculated by Hedges' g with 95% CIs, p values, z-scores, heterogeneity and forest plots) are in Appendix A, supplementary results – individual studies meta-analyses.

Primary outcomes: kinematic measures

All KINOs demonstrated statistically significant responsiveness, except for trunk ROM. We found statistically significant moderate pooled effect estimates demonstrating improvement in post-test scores overall for smoothness (Fig. A1), movement duration (Fig. A2; improvement expressed as reduction in time scores), peak velocity (Fig. A3), shoulder ROM (Fig. A4), control strategy (Fig. A5) and movement efficiency (Fig. A6; expressed as a reduction in the measure, hence a negative value). Elbow ROM (Fig. A7) demonstrated a statistically significant small to moderate improvement in pooled effect estimate post-test scores, but trunk ROM (Fig. A8) did not demonstrate a statistically significant improvement in pooled effect estimate post-test scores.

Clinical outcomes measures

All CLIOs demonstrated statistically significant responsiveness. The MAL questionnaire subscales amount of use (AOU) (Fig. A9) and quality of movement (QOM) (Fig. A10) demonstrated large statistically significant effect sizes for post-test measures. FMA-UE (Fig. A11), WMFT Time subscale (Fig. A12; expressed as a reduction to indicate improvement) and WMFT Quality (Fig. A13) score variables demonstrated statistically significant moderate improvement of the pooled effect estimate.

ARAT and BBT were reported in only one trial, so meta-analysis was inappropriate. BBT reported in a TRT subgroup [57] (n=15) demonstrated significant improvement (p<0.01), but no CI was provided. Improvements of the activity measure ARAT was reported [53] in CIMT+TRT (n=20) and CIMT (n=19) groups (p<0.05), but no CI was provided. No studies reported the use of the CAHAI to measure UL functional capacity in conjunction with KINO analysis post-stroke.

Comparison of responsiveness of KINOs and CLIOs

Comparison of the pooled effect estimates of SMDs calculated by Hedges' g (95% CI) for individual KINOs and CLIOs are shown in Figure 3. Comparison of pooled effect estimates enabled identification of the strength of change after interventions.

No statistical difference of responsiveness can be assumed between measures, except for MAL AOU and QOM over KINO movement efficiency, trunk ROM, shoulder ROM, control strategy and peak velocity due to absence of overlapping CIs.

Smoothness was the most responsive KINO, with the highest moderate estimated effect size, whereas the FMA-UE clinical measure of UL structure and function demonstrated similar responsiveness. In descending order of responsiveness effect sizes, moderate effect sizes were demonstrated across WMFT Time measure, movement duration, WMFT Quality, peak velocity, and shoulder ROM, followed by movement efficiency, control strategy and elbow ROM. Finally, trunk ROM did not demonstrate significant responsiveness.

Sensitivity analysis

The sensitivity analysis conducted with different correlation values of pre–post measures for calculation of the between-outcomes summative effect size (i.e., responsiveness for KINOs and CLIOs) revealed that the analyses with r = 0.5 were robust. Changing the correlations had a negligible impact on the results (Table A1 in Appendix A, supplementary results – sensitivity analysis).

Discussion

The main finding of this review was the reporting and comparison of responsiveness of UL stroke outcome measures as determined by treatment effect sizes of pre-test/post-test measurements, after

CIMT, TR, CIMT+TR and BAT rehabilitation. Thirteen KINO (n=8) and CLIO (n=5) outcome measures were identified, which allowed for a comparative overview of meta-analysis results. All measures demonstrated significant statistical improvements, with the exception of trunk ROM. Responsiveness of kinematic variables was identified for use in clinical trials and practice, filling a gap identified in the literature [22]. Information regarding psychometric properties of KINO measures will assist in the selection of clinically relevant kinematic variables appropriate for intervention studies and practice [22].

Comparison of responsiveness between outcomes measures

MAL activity measure was present among all 9 studies reporting CIMT [5,46,48–52,56]. The responsiveness of MAL AOU and QOM was greater than peak velocity, shoulder ROM, movement efficiency, control strategy, elbow ROM and trunk ROM. Comparison of effect sizes of kinematic and clinical outcomes measures demonstrated homogenous responsiveness as demonstrated by overlap of 95% CIs [42]. Contrary to previous studies of similar populations, responsiveness of KINOs was not greater than CLIOs after CIMT, TR, CIMT+TR and BAT in patients with chronic stroke [13,22,26] and no advantage of peak velocity and movement duration was demonstrated over WMFT Time [10,18]. This discrepancy can be explained by the greater statistical power of the meta-analysis as compared with individual studies.

Consideration of subjective measures

The subjective dimension of the MAL questionnaire should be highlighted in terms of the influence of treatment effects leading to effect sizes larger than the more objective measures, based on actual performance [7,58,59]. This finding illustrates potential overestimation by subjective measures on treatment effects contributing to the risk of bias in outcome assessment [6,11,29]. The benefit of interventions judged by a subjective questionnaire (i.e., MAL scores) appeared greater than that judged by objective outcomes measures (i.e., kinematic data or scales with standardized rating by a therapist). This finding raises questions for future research to compare MAL to quantitative ecological observational measures [1,6], highlighting the importance of using MAL combined with outcome measures of various ICF levels to ensure objective and comprehensive evaluation of UL impairment on activities of daily living [6,7,10,11].

Trends and combination of measure selection

CLIOs: FMA-UE was the most reported measure combined with kinematic analysis, which is consistent with findings supporting its use for assessing body function and structure in conjunction with activity ICF measures, as a predominant consensus post-stroke [6,10,13]. Trends of ICF activity-measure use across trials varied. Despite emerging international consensus supporting the use of ARAT [13], it was reported in only one study [52]. Trends of WMFT and BBT use, reported in 2 and 1 studies, respectively, are consistent with reviews identifying the highest frequency of use among these measures with ARAT and MAL [6,10]. Despite well-established psychometric properties in post-stroke rehabilitation [10], CAHAI was not reported in conjunction with KINOs.

KINOs: Movement duration, efficiency and control strategy were the most reported kinematic variables, followed by smoothness, elbow ROM, peak velocity, shoulder ROM and trunk displacement. KINO selection was consistent with findings among stroke populations across 93 intervention, comparative and longitudinal trials [22]. Our findings are consistent with recommendations for the use of measures in conjunction with ICF levels [6,7,10].

Although the literature recommends measure classification according to ICF levels, time since stroke, measurement type and psychometric properties, the responsiveness of KINO variables remains understudied [6,13,22]. Increased use of kinematic analysis in UL evaluation highlights a need to standardise analysis methods and the identification of best-fit variables for specific interventions and prediction of recovery, encouraging their potential use across clinical practice [10,15,19,22].

Limitations

The lack of availability of data from studies, despite contact with authors, did not provide sufficient information for potential imputation of correlation values appropriate to repeated-measures analysis methodology. Restriction of study selection to English and French publications may have contributed to a language bias [29,30]. Despite application of a duplicate publication decision tree [30], multiple publication bias may have influenced results. Even if predefined inclusion criteria and RCT design selection contributed to comparability and synthesis of results [29], evidence can only be generalised to CIMT, TRT and BAT interventions using 3D optoelectronic movement analysis systems [22]. Variation of data acquisition conditions across studies, such as number of trials, task analysed (e.g., reaching with or without grasping) and/or sampling rate, may have affected the comparability of effect sizes [9,22,23]. Studies had low or unclear bias, with the exception of 2 studies presenting high risk of bias in conjunction with small sample sizes. The conclusions of this work cannot be extrapolated to kinematic data obtained from virtual, robotic or haptic devices, which were not retained in our analysis in order to limit the heterogeneity of the data.

Conclusions

This study has highlighted current trends of measure selection for evaluating UL function post-stroke after CIMT, BAT and TRT. Findings support the combined use of kinematic analysis and clinical outcome measures aligned with ICF levels for assessment. We identified the use of the FMA-UE measure for assessing ICF body function level, which supports international consensus. Variations in the use of ICF activity level measures indicate that further consensus is needed regarding their use across trials. Furthermore, the effect of objective (i.e., kinematic data or scales with standardized rating by a therapist) or subjective (i.e., self-reported questionnaire) nature of outcomes measures on treatment effects should be considered in trials, because findings confirm that subjective measures demonstrated greater perceived benefits. Kinematic variables did not have higher levels of responsiveness over objective clinical outcome measures in adults with chronic stroke. However, unlike the CLIOs, they have the advantage of characterizing the quality and structure of the movement, leading to a better understanding of the underlying neural mechanisms of functional improvements [9].

Further research should include various study designs, rehabilitation techniques and times post-stroke to enable further investigation for use of standardised kinematic variables. The development of wearable, inexpensive and easy-to-use kinematic analysis devices could promote their dissemination

[23]. Consensus will contribute to the identification of clinically significant changes contributing to evidence-based practice. Future research of kinematic-variable responsiveness for stroke rehabilitation should highlight the potential use of combined evaluative methods, demonstrating psychometric qualities for composite scores of UL function combined with ICF components.

Acknowledgements. The authors acknowledge the authors of the included studies in the review who provided us with additional information enabling data analysis.

Conflict of interest. None declared.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Figure legends

Figure 1: PRISMA flow diagram of study selection.

Figure 2: Risk of bias of studies assessed with the Cochrane Collaboration Risk of Bias tool.

Figure 3: Responsiveness assessed by comparing meta-analysis pooled effect estimates of standardised mean difference obtained with a correlation coefficient of 0.5 between pre- and post-test paired data for individual KINOs and CLIOs (CI, confidence interval; CLIO, clinical outcome measures; FMA-UE, Fugl-Meyer Assessment - Upper Extremity; KINO, kinematic outcome measures; MAL, Motor Activity Log; AOU, amount of use; QOM, quality of movement; ROM, range of motion; WMFT, Wolf Motor Function Test).

References

[1] Barden HLH, Baguley IJ, Nott MT, Heard R, Chapparo C. Measuring task performance after acquired brain injury: Construct and concurrent validity of “Upper Limb Performance Analysis.” Brain Inj 2015:1–9. https://doi.org/10.3109/02699052.2015.1028446.

[2] Jørgensen HS, Nakayama H, Raaschou HO, Vive-Larsen J, Støier M, Olsen TS. Outcome and time course of recovery in stroke. Part I: Outcome. The Copenhagen Stroke Study. Arch Phys Med Rehabil 1995;76:399–405. https://doi.org/10.1016/S0003-9993(95)80567-2.

[3] Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. The Lancet 2011;377:1693–702. https://doi.org/10.1016/S0140-6736(11)60325-5.

[4] Kwakkel G. Towards integrative neurorehabilitation science. Physiother Res Int 2009;14:137–46 10p. https://doi.org/10.1002/pri.446.

[5] Lima RCM, Michaelsen SM, Nascimento LR, Polese JC, Pereira ND, Teixeira-Salmela LF. Addition of trunk restraint to home-based modified constraint-induced movement therapy does not bring additional benefits in chronic stroke individuals with mild and moderate upper limb impairments: A pilot randomized controlled trial. NeuroRehabilitation 2014;35:391–404. https://doi.org/10.3233/NRE-141130.

[6] Santisteban L, Térémetz M, Bleton J-P, Baron J-C, Maier MA, Lindberg PG. Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. PloS One 2016;11:e0154792. https://doi.org/10.1371/journal.pone.0154792.

[7] Lemmens RJM, Timmermans AAA, Janssen-Potten YJM, Smeets RJEM, Seelen HAM. Valid and reliable instruments for arm-hand assessment at ICF activity level in persons with hemiplegia: a systematic review. BMC Neurol 2012;12:21. https://doi.org/10.1186/1471-2377-12-21.

[8] Levin MF, Kleim JA, Wolf SL. What do motor “recovery” and “compensation” mean in patients following stroke? Neurorehabil Neural Repair 2009;23:313–9. https://doi.org/10.1177/1545968308328727.

[9] Demers M, Levin MF. Do Activity Level Outcome Measures Commonly Used in Neurological Practice Assess Upper-Limb Movement Quality? Neurorehabil Neural Repair 2017;31:623–37. https://doi.org/10.1177/1545968317714576.

[10] Alt Murphy M, Resteghini C, Feys P, Lamers I. An overview of systematic reviews on upper extremity outcome measures after stroke. BMC Neurol 2015;15:29. https://doi.org/http://dx.doi.org/10.1186/s12883-015-0292-6.

[11] Ashford S, Slade M, Malaprade F, Turner-Stokes L. Evaluation of functional outcome measures for the hemiparetic upper limb: a systematic review. J Rehabil Med 2008;40:787–95. https://doi.org/10.2340/16501977-0276.

[12] Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. Lancet Neurol 2007;6:1094–105. https://doi.org/10.1016/S1474-4422(07)70290-9.

[13] Kwakkel G, Lannin NA, Borschmann K, English C, Ali M, Churilov L, et al. Standardized measurement of sensorimotor recovery in stroke trials: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2017;12:451–61. https://doi.org/10.1177/1747493017711813.

[14] Lang CE, Bland MD, Bailey RR, Schaefer SY, Birkenmeier RL. Assessment of upper extremity impairment, function, and activity after stroke: foundations for clinical decision making. J Hand Ther Off J Am Soc Hand Ther 2013;26:104–14;quiz 115. https://doi.org/10.1016/j.jht.2012.06.005.

[15] Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF activity. Disabil Rehabil 2005;27:315–40. https://doi.org/10.1080/09638280400008545.

[16] Alt Murphy M, Willén C, Sunnerhagen KS. Responsiveness of upper extremity kinematic measures and clinical improvement during the first three months after stroke. Neurorehabil Neural Repair 2013;27:844–53. https://doi.org/10.1177/1545968313491008.

[17] Veerbeek JM, van Wegen E, van Peppen R, van der Wees PJ, Hendriks E, Rietberg M, et al. What Is the Evidence for Physical Therapy Poststroke? A Systematic Review and Meta-Analysis. PLoS ONE 2014;9. https://doi.org/10.1371/journal.pone.0087987.

[18] Rohafza M, Fluet GG, Qiu Q, Adamovich S. Correlation of reaching and grasping kinematics and clinical measures of upper extremity function in persons with stroke related hemiplegia. Conf

Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf 2014;2014:3610–3. https://doi.org/10.1109/EMBC.2014.6944404.

[19] Li K, Lin K, Chen C, Liing R, Wu C, Chang W. Concurrent and Predictive Validity of Arm Kinematics With and Without a Trunk Restraint During a Reaching Task in Individuals With Stroke. Arch Phys Med Rehabil 2015;96:1666–75. https://doi.org/10.1016/j.apmr.2015.04.013.

[20] Bernhardt J, Hayward KS, Kwakkel G, Ward NS, Wolf SL, Borschmann K, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: The Stroke Recovery and Rehabilitation Roundtable taskforce. Int J Stroke Off J Int Stroke Soc 2017;12:444–50. https://doi.org/10.1177/1747493017711816.

[21] Térémetz M, Colle F, Hamdoun S, Maier MA, Lindberg PG. A novel method for the quantification of key components of manual dexterity after stroke. J NeuroEngineering Rehabil 2015;12. https://doi.org/10.1186/s12984-015-0054-0.

[22] Murphy MA, Häger CK. Kinematic analysis of the upper extremity after stroke – how far have we reached and what have we grasped? Phys Ther Rev 2015. https://doi.org/10.1179/1743288X15Y.0000000002.

[23] Kwakkel G, van Wegen E, Burridge JH, Winstein CJ, Alt Murphy M, Levin MF, et al. Standardized measurement of quality of upper limb movement after stroke: Consensus-based core recommendations from the Second Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2019;14:783–91. https://doi.org/10.1177/1747493019873519.

[24] Stinear C. Prediction of recovery of motor function after stroke. Lancet Neurol 2010;9:1228–32. https://doi.org/10.1016/S1474-4422(10)70247-7.

[25] Edwards DF, Lang CE, Wagner JM, Birkenmeier R, Dromerick AW. An evaluation of the Wolf Motor Function Test in motor trials early after stroke. Arch Phys Med Rehabil 2012;93:660–8. https://doi.org/10.1016/j.apmr.2011.10.005.

[26] Caimmi M, Carda S, Giovanzana C, Maini ES, Sabatini AM, Smania N, et al. Using kinematic analysis to evaluate constraint-induced movement therapy in chronic stroke patients. Neurorehabil Neural Repair 2008;22:31–9. https://doi.org/10.1177/1545968307302923.

[27] Hebert JS, Lewicke J, Williams TR, Vette AH. Normative data for modified Box and Blocks test measuring upper-limb function via motion capture. J Rehabil Res Dev 2014;51:918–32. https://doi.org/10.1682/JRRD.2013.10.0228.

[28] Alt Murphy M, Willén C, Sunnerhagen KS. Responsiveness of upper extremity kinematic measures and clinical improvement during the first three months after stroke. Neurorehabil Neural Repair 2013;27:844–53. https://doi.org/10.1177/1545968313491008.

[29] Higgins J, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration; 2011.

[30] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700. https://doi.org/10.1016/j.jclinepi.2009.06.006.

[31] De los Reyes-Guzmán A, Dimbwadyo-Terrer I, Trincado-Alonso F, Monasterio-Huelin F, Torricelli D, Gil-Agudo A. Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: A review. Clin Biomech 2014;29:719–27. https://doi.org/10.1016/j.clinbiomech.2014.06.013.

[32] Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459–68. https://doi.org/10.1016/S0895-4356(99)00206-1.

[33] Elamin MB, Flynn DN, Bassler D, Briel M, Alonso-Coello P, Karanicolas PJ, et al. Choice of data extraction tools for systematic reviews depends on resources and review complexity. J Clin Epidemiol

2009;62:506–10. https://doi.org/10.1016/j.jclinepi.2008.10.016.

[34] Borenstein M, Hedges L, Higgins J, Rothstein H. Introduction to Meta-Analysis. John Wiley & Sons Ltd; 2009.

[35] Peters JL, Mengersen KL. Meta-analysis of repeated measures study designs. J Eval Clin Pract 2008;14:941–50. https://doi.org/10.1111/j.1365-2753.2008.01010.x.

[36] Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60. https://doi.org/10.1136/bmj.327.7414.557.

[37] Borenstein M, Higgins JPT, Hedges LV, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Res Synth Methods 2017;8:5–18. https://doi.org/10.1002/jrsm.1230.

[38] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010;1:97–111. https://doi.org/10.1002/jrsm.12.

[39] Akobeng AK. Understanding systematic reviews and meta-analysis. Arch Dis Child 2005;90:845–8. https://doi.org/10.1136/adc.2004.058230.

[40] Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. J Stat Softw 2010;36:1–48. https://doi.org/10.18637/jss.v036.i03.

[41] Del Re AC. A practical tutorial on conducting meta-analysis in R. Quant Methods Psychol 2015;11:37–50. https://doi.org/10.20982/tqmp.11.1.p037.

[42] Knol MJ, Pestman WR, Grobbee DE. The (mis)use of overlap of confidence intervals to assess effect modification. Eur J Epidemiol 2011;26:253–4. https://doi.org/10.1007/s10654-011-9563-8.

[43] Thielman G. Insights into upper limb kinematics and trunk control one year after task-related training in chronic post-stroke individuals. J Hand Ther 2013;26:156–61.

https://doi.org/10.1016/j.jht.2012.12.003.

[44] Lang CE, Wagner JM, Dromerick AW, Edwards DF. Measurement of Upper-Extremity Function Early After Stroke: Properties of the Action Research Arm Test. Arch Phys Med Rehabil 2006;87:1605–10. https://doi.org/10.1016/j.apmr.2006.09.003.

[45] Thielman G, Kaminski T, Gentile AM. Rehabilitation of reaching after stroke: comparing 2 training protocols utilizing trunk restraint. Neurorehabil Neural Repair 2008;22:697–705. https://doi.org/10.1177/1545968308315998.

[46] Wu C, Chen C, Tang SF, Lin K, Huang Y. Kinematic and Clinical Analyses of Upper-Extremity Movements After Constraint-Induced Movement Therapy in Patients With Stroke: A Randomized Controlled Trial. Arch Phys Med Rehabil 2007;88:964–70. https://doi.org/10.1016/j.apmr.2007.05.012.

[47] Pang MY, Harris JE, Eng JJ. A Community-Based Upper-Extremity Group Exercise Program Improves Motor Function and Performance of Functional Activities in Chronic Stroke: A Randomized Controlled Trial. Arch Phys Med Rehabil 2006;87:1–9. https://doi.org/10.1016/j.apmr.2005.08.113.

[48] Lin K-C, Wu C-Y, Wei T-H, Lee C-Y, Liu J-S. Effects of modified constraint-induced movement therapy on reach-to-grasp movements and functional performance after chronic stroke: a randomized controlled study. Clin Rehabil 2007;21:1075–86. https://doi.org/10.1177/0269215507079843.

[49] Woodbury ML, Howland DR, McGuirk TE, Davis SB, Senesac CR, Kautz S, et al. Effects of Trunk Restraint Combined With Intensive Task Practice on Poststroke Upper Extremity Reach and Function: A Pilot Study. Neurorehabil Neural Repair 2009;23:78–91. https://doi.org/10.1177/1545968308318836.

[50] Wu C, Lin K, Chen H, Chen I -hsue., Hong W. Effects of modified constraint-induced movement therapy on movement kinematics and daily function in patients with stroke: a kinematic

study of motor control mechanisms. Neurorehabil Neural Repair 2007;21:460–6. https://doi.org/10.1177/1545968307303411.

[51] Wu C, Chuang L, Lin K, Chen H, Tsay P. Randomized trial of distributed constraint-induced therapy versus bilateral arm training for the rehabilitation of upper-limb motor control and function after stroke. Neurorehabil Neural Repair 2011;25:130–9. https://doi.org/10.1177/1545968310380686.

[52] Wu C, Chen Y, Chen H, Lin K, Yeh I -lin. Pilot trial of distributed constraint-induced therapy with trunk restraint to improve poststroke reach to grasp and trunk kinematics. Neurorehabil Neural Repair 2012;26:247–55. https://doi.org/10.1177/1545968311415862.

[53] Wu C, Chen Y, Lin K, Chao C, Chen Y. Constraint-Induced Therapy With Trunk Restraint for Improving Functional Outcomes and Trunk-Arm Control After Stroke: A Randomized Controlled Trial. Phys Ther 2012;92:483–92 10p. https://doi.org/10.2522/ptj.20110213.

[54] Cacho R de O, Cacho EWA, Ortolan RL, Cliquet A, Borges G. Trunk Restraint Therapy: The Continuous Use of the Harness Could Promote Feedback Dependence in Poststroke Patients A Randomized Trial. Medicine (Baltimore) 2015;94:e641. https://doi.org/10.1097/MD.0000000000000641.

[55] Michaelsen SM, Dannenbaum R, Levin MF. Task-specific training with trunk restraint on arm recovery in stroke - Randomized control trial. Stroke 2006;37:186–92. https://doi.org/10.1161/01.STR.0000196940.20446.c9.

[56] Lin K, Chen Y, Chen C, Wu C, Chang Y. The effects of bilateral arm training on motor control and functional performance in chronic stroke: a randomized controlled study. Neurorehabil Neural Repair 2010;24:42–51. https://doi.org/10.1177/1545968309345268.

[57] Michaelsen SM, Luta A, Roby-Brami A, Levin MF. Effect of Trunk Restraint on the Recovery of Reaching Movements in Hemiparetic Patients. Stroke 2001;32:1875–83.

https://doi.org/10.1161/01.STR.32.8.1875.

[58] Taub E, Crago JE, Uswatte G. Constraint-induced movement therapy: A new approach to treatment in physical rehabilitation. Rehabil Psychol 1998;43:152–70. https://doi.org/10.1037/0090-5550.43.2.152.

[59] Uswatte G, Taub E, Morris D, Light K, Thompson PA. The Motor Activity Log-28 Assessing daily use of the hemiparetic arm after stroke. Neurology 2006;67:1189–94. https://doi.org/10.1212/01.wnl.0000238164.90657.c2.