Neurorehabil Neural Repair. Author manuscript; available in PMC 2015 July 01.

Published in final edited form as:
Neurorehabil Neural Repair. 2013 ; 27(3): 240–250. doi:10.1177/1545968312461719.

Functional recovery following stroke: Capturing changes in upper extremity function

Lisa A Simpson, BSc and
Graduate Program in Rehabilitation Sciences, University of British Columbia, Rehabilitation Research Lab, GF Strong Rehabilitation Centre, 4255 Laurel St., Vancouver, BC, Canada, V5Z 2G9

Janice J. Eng, PhD, PT/OT [Professor]
Department of Physical Therapy, University of British Columbia, 212-2177 Wesbrook Mall, Vancouver, BC, Canada, V6T 1Z3, Tel: 604-714-4105 Fax: 604-714-4168

Abstract

Background and Purpose—Augmenting changes in recovery is core to the rehabilitation process following a stroke. Hence, it is essential that outcome measures are able to detect change as it occurs; a property known as responsiveness. This paper critically reviewed the responsiveness of functional outcome measures following stroke, specifically examining tools that captured upper extremity functional recovery.

Methods—A systematic search of the literature was undertaken to identify articles providing responsiveness data for three types of change (observed, detectable, important).

Results—Data from 68 articles for 14 upper extremity functional outcome measures were retrieved. Larger percent changes were required to be considered important when obtained through anchor-based methods (eg. based on patient opinion or comparative measure) compared to distribution methods (eg. statistical estimates). Larger percent changes were required to surpass the measurement error for patient-perceived functional measures (eg. Motor Activity Log) compared to lab-based performance measures (eg. Action Research Arm Test). The majority of rehabilitation interventions have similarly sized effects on patient-perceived upper extremity function versus lab-based upper extremity function.

Conclusions—The magnitude of important change or change that surpasses measurement error can vary substantially depending on the method of calculation. Rehabilitation treatments can affect patient perceptions of functional change as effectively as lab-based functional measures; however higher sample sizes may be required to account for the larger measurement error associated with patient-perceived functional measures.

Keywords

stroke; systematic review; recovery of function; responsiveness; treatment outcome; rehabilitation; upper extremity

Correspondence to: Janice J. Eng.

INTRODUCTION

Functional recovery following stroke is complex with wide variation in natural recovery and response to treatment across individuals. Optimizing or augmenting changes in recovery is core to the rehabilitation process following a stroke. Hence, it is essential that outcome measures are able to detect change as it occurs; a property known as responsiveness.[1,2]

Detecting change over time or from an intervention is one of the most critical requisites of an outcome measure; it is necessary information for selecting the best instrument for practice or research and for determining sample size for clinical studies. Furthermore, there is growing recognition that traditional reporting of statistical significance tests and effects sizes should be accompanied with methods for determining meaningful or important change.[3]

Beaton et al.[1] argue that responsiveness is a context-specific characteristic that is influenced by factors such as the specific sample, treatment, and the type of change captured by an instrument. The authors outline three major types of change in their responsiveness taxonomy. They are: 1) observed change, 2) important change and 3) detectable change. Figure 1 depicts the inter-relationship between the three types of change and the common metrics used to quantify them. Observed change is the amount of change observed in a population in which change is expected to occur (ie. after a treatment of known efficacy or a specific period within the natural recovery pattern). Traditional methods have captured this type of change with an effect size.[1] Important change is the observed change estimated to be meaningful and is often quantified as a minimally clinically important difference value (MCID). For example, important change may reflect the value that patients, clinicians or society places on the recovery (or partial recovery) of a task, like the ability to bring a spoon to the mouth. Finally, detectable change takes into consideration the measurement error associated with a tool and is often quantified as a minimal detectable change score (MDC) or limit of agreement (LOA). Calculation of the MDC value associated with a measure varies depending on the confidence level selected. The most common confidence levels selected are 95% and 90% and are denoted by the subscripts MDC_{95} and MDC_{90}. Confidence that true functional change has been observed in a clinical study is increased when the observed change is equal to or surpasses a measure's detectable change values.[1]

The purpose of this paper is to synthesize and critically review the research evidence that captures responsiveness as defined by three types of change (observed, important, detectable). This systematic review provides an understanding of the responsiveness of outcome measures used in stroke research, specifically within the context of upper extremity (UE) functional recovery.

METHODS

Literature Search

This review targeted articles that capture three types of change (observed, important, detectable) using UE functional outcome measures following a stroke. Articles were identified using a systematic search of electronic databases (MEDLINE, EMBASE,

CINAHL, PsycINFO, Cochrane CENTRAL) from database inception through March 2012. The following keywords were used: stroke, cerebrovascular accident or hemiplegia or hemiparesis combined with upper extremity, function or activities of daily living and responsiveness, reliability, psychometrics, “minimally clinically important change or MCID”, “standard error of measurement or SEM”, “minimal detectable change or MDC”, “standardized response mean or SRM”, effect size or outcome measurement. All terms were mapped onto subject headings. Articles were limited to the English language, human subjects, and adults. A hand search of reference lists from reviews and the grey literature (eg. StrokEngine Assess) was also conducted to ensure a thorough search.[4]

Inclusion Criteria

To be included in this review, articles had to meet the following inclusion criteria: 1) provided a responsiveness index or sufficient information to calculate responsiveness index (eg. minimally detectable change or effect size); 2) utilized a sample of individuals with stroke; 3) used an outcome measure that assessed affected UE ability and included ≥50% functional activities (to ensure scale reflected International Classification of Functioning, Disability and Health activity domain) and 4) published in a peer-reviewed journal. Exclusion criteria consisted of: 1) conference proceedings or abstracts; 2) pre-post studies or randomized control trials (RCTs) which calculated a responsiveness index without utilizing the control group information to minimize bias for inflation of the effect sizes;[5] 3) studies with a sample size less than ten; 4) articles that utilized measures that captured UE function as a single action (eg. Box and Blocks test, peg test). Single actions were not considered representative of the many actions involved in UE function. RCTs were excluded if they: 1) utilized <2 UE functional outcome measures or 2) found no significant effects for UE function.

Data extraction and organization

Responsiveness indices were extracted or calculated from the included articles and subsequently organized into the three categories of change outlined by Beaton et al.[1]: 1) observed change, 2) important change and 3) detectable change (Figure 1). Observed change was further subdivided into 1) change over natural recovery (categorized into <3 months post injury and ≥ 3 months post injury) and 2) change in response to an effective treatment. Effect sizes were calculated based on the change score divided by the baseline standard deviation and minimally detectable change values ($MDC_{90,}$ MDC_{95}) were calculated based on the test-retest coefficient (ICC) and the baseline standard deviation.[6] In addition, MDC_{95}%, which are independent of measurement units, were calculated to compare minimal detectable change values across measures.[7] We present the MDC_{95}% using the two methods commonly cited in the rehabilitation literature: 1) MDC_{95}/maximum score for the scale (eg. 57 points for ARAT) and 2) MDC_{95}/baseline mean of the sample.

RCTs that utilized at least two different functional outcome measures were used to estimate observed change in response to a treatment of known efficacy. This allowed for comparison of observed change across different measures within the same study by controlling for variation in treatment and sample characteristics.[8] To ensure we were capturing observed change in response to an effective treatment, only RCTs that found a significant effect for at

least one of the UE functional measures were included. Scatterplots of effect sizes were generated for a visual representation of one outcome measures' relative ability to capture change compared to another measure in response to the same treatment with the same sample. When only median and range scores were provided, mean and standard deviation values were estimated using the method suggested by Hozo et al.[9] in order to calculate an effect size.

RESULTS

The search strategy yielded 1770 titles of which 68 met the inclusion criteria (Figure 2). The articles provided responsiveness data for 14 functional outcome measures.[10–77] Supplementary Table 1 displays the number of articles that provided responsiveness estimates for each measure. The Action Research Arm Test (ARAT), Motor Activity Log (MAL), Wolf Motor Function test (Wolf) and Stroke Impact Scale (SIS) were the four most frequently used measures among the included articles.

Observed change

Observed Change over natural recovery—Effect sizes that captured change over time of participants in standard care were extracted or calculated from 25 studies.[10–34] Fourteen studies followed participants up to 3 months post stroke and ten studies followed participants up to 6 months post stroke. As only one study followed participants past 6 months post stroke, we compared the effect sizes for participants less than 3 months versus those greater or equal to 3 months post injury. The majority of studies had observation periods from 2 to 5 months.

Effect sizes calculated at a baseline of 1–3 months post injury were larger and showed greater variance than effect sizes calculated at a baseline ≥ 3 months post injury (Figure 3). This finding was evident when the duration of follow-up was similar between the two groups. Two studies[17,25] considered the effect of stroke severity on observed changes over recovery; effect sizes were 2.0–2.6 times larger for individuals with less severe impairments at 1–2 months post-stroke at study baseline. Three studies[13,31,33] calculated effect sizes using two methods: population effect size (based on the change score divided by baseline standard deviation) and standardized response mean (based on the change score divided by the change score standard deviation) for the same measure. In all three studies, the effect sizes calculated as a standardized response mean were larger than when calculated as a population effect size (Supplementary Table 2).

Observed change in response to a treatment of known efficacy—A total of 28 RCTs utilized more than one UE functional outcome measure and obtained a significant effect for at least one of these measures.[35–62] These 28 RCTs were used to examine observed change of different measures in response to a treatment of known efficacy within the same study. Constraint-Induced Movement Therapy (CIMT) was the most frequently studied intervention among the included RCTs (50%). Also, the majority of these studies utilized a lab-based performance measure (eg. ARAT, Wolf), in addition to a measure that captured perceived function in one's own environment (eg. MAL, SIS) (n=23). The most

common outcome measures used together were the MAL with the ARAT (used in 7 studies) [35–41] or MAL with the Wolf (used in 10 studies).[42–51]

The majority of effect sizes were close to a 1:1 relationship between the lab-based versus patient-perceived functional measures within the same study (Figure 4). Effect sizes from studies which investigated CIMT[36–38,46,47] and one study that examined the effect of mirror therapy for individuals with Complex Regional Pain Syndrome[43] did not demonstrate this 1:1 relationship. Effect sizes for the perceived effect (MAL) were 1.66.2 times larger than the functional changes (measured by ARAT or Wolf) in these studies.

Important Change

Five studies established important change (MCID) values for six functional outcome measures (See Table 1).[63–67] All but one study calculated MCID values from CIMT trials.[63,64,66,67] The other study investigated change following robotic therapy.[65] Four studies[63,65–67] utilized individuals with chronic stroke and one study sample consisted of individuals with acute stroke.[64]

Three studies[65–67] used a combination of anchor and distribution based techniques whereas two studies[63,64] used only anchor based methods to define important change. Anchor-based methods compare the change scores on the measure of interest to a comparative measure or 'anchor' of important change.[78] The following anchors were used in the five studies: 1) a predetermined level on a global rating scale in which participants were asked to rate their perception of functional change;[64] 2) a predetermined level of recovery on the Stroke Impact Scale (SIS) global recovery item[63,65,66] and 3) a change score of 6–10 points on the upper extremity portion of the Fugl-Meyer assessment.[67] Distribution based methods determine important change based on the statistical distribution of the results.[78] Three studies used 0.2 times the standard deviation of the sample (i.e. effect size of 0.2) to determine MCID values.[65–67]

Important change values displayed large variation with values spanning from 1.1%–30% of the tests' maximum scores. The largest values were observed in the study that utilized a sample of individuals in the acute stage post stroke.[64] In addition, MCID values calculated using statistically-derived distribution based methods were substantially smaller (15–88% less) than the values determined by anchor based methods (Table 1). For example, the MCID for the SIS-hand ranged from a value of 5.8 using an effect size method (distribution method) to 17.8 using perceived amount of recovery on the global recovery question (participant-perceived anchor-based method) for the same sample of individuals.[66]

Three studies were not able to establish MCID values due to non-significant relationships between the global recovery/rating scales and functional changes using the MAL,[63,76] Wolf (time component),[63,64] and accelerometry[64] (Table 1).

Detectable Change

Values needed to surpass measurement error, which are considered to represent true functional change (SEM, MDC_{90}, MDC_{95} and LOA values), were extracted or calculated from 16 studies for 9 measures.[12,14,15,27,66–77] All studies used a test-retest methodology in

which time frames in between assessments ranged from 1 day to 2 weeks. Nine studies utilized individuals with subacute stroke[14,15,27,68–70,72–74] and seven studies used samples with chronic stroke.[12,66,67,71,75–77] Nine studies utilized a subsample of individuals from CIMT/forced use trials[66–72,76,77] of which four were from the EXCITE trial.[68–70,72]

Minimal detectable change values at the 90% and 95% confidence levels ranged from 1.0% and 1.2% of the maximum score for the AMAT-time subscale to 21.9% and 25.9% of the maximum score for the SIS-hand respectively. Relative to their sample means, minimal detectable change values at the 90% and 95% confidence levels ranged from 11.5% and 13.7% for the Wolf-functional ability subscale to 72.5% and 86.7% for the MAL-amount of use subscale respectively. Detailed SEM, MDC_{90} and MDC_{95} estimates and their respective relative percent values (ie. relative to the sample mean or relative to the scale maximum score) can be found in Supplementary Table 3 available online. Among the measures with multiple estimates, the values needed to surpass measurement error for tools that capture patient-perceived function (eg. MAL, SIS) were larger than lab-based performance measures (eg. Wolf, CAHAI) (Figure 5). This observation was present whether the minimal detectable change values were considered relative to the scale maximums or the sample means.

Also, the values needed to surpass measurement error for patient-perceived performance measures were larger or on par with important change values. For instance, minimal detectable change and important change values for the SIS were: 17.1–21.9 (MDC_{90}) and 20.4–25.9 (MDC_{95}) versus 5.8 (distribution-based MCID) or 17.8 (anchor-based MCID). Minimal detectable change and important change values for the MAL were: 0.56–1.06 (MDC_{90}) and 0.67–1.27 (MDC_{95}) versus 1.0–1.1 (anchor-based MCID).

DISCUSSION

Examining the measurement of change in UE function served as a framework for understanding the measurement of functional recovery following stroke and revealed several novel findings related to the ability of outcome measures to capture change as it occurs.

Observed Change

For equal duration of follow up, the effect sizes due to natural recovery calculated at 1–3 months post stroke were substantially larger than those calculated at 3 months or later post stroke. The observed differences in effect sizes between these time phases likely reflect the higher degree of neuroplasticity that have been documented early after stroke.[79] The effect sizes obtained in the RCTs using a population at >6 months post stroke ranged from 0.05 to 4.28 demonstrating that individuals are still capable of change at later time periods post stroke when receiving treatment. Of importance, our collective data demonstrate that rehabilitation treatments can affect patient perceptions of functional change as effectively as lab-based functional measures. In fact, CIMT is an exemplary treatment model where patient perceptions of change (MAL) were 1.6–6.2 times larger than effect sizes obtained with lab-based functional performance measures (ARAT or Wolf). A likely explanation is that CIMT was specifically designed to overcome learned non-use, thereby targeting functional change in one's own environment. Additionally, the higher relative effect sizes observed in the MAL following CIMT could also reflect patient bias. This bias could be present due to the large

investment of time and effort required from the CIMT program. Finally, it should be noted that our observations concerning the relative ability of treatments to affect lab-based versus patient-perceived function may be limited to studies able to capture statistically significant effects.

Real time accelerometry monitoring of UE activity in the home and community is a promising technology to objectively capture function in one's own environment without the necessity of self-report.[80] However, only one study provided an estimate of true or important change for accelerometry measures. More studies are needed to better understand the usefulness of real time activity monitors for capturing the effectiveness of UE rehabilitation interventions.

The influence of the method used to calculate effect sizes (ie. methods based on baseline standard deviation versus standard deviation of the change scores) was also revealed. Methods based on the change score standard deviation (ie. standardized response mean) produced estimates that were up to 1.2–1.9 times larger than those calculated based on the baseline standard deviation (ie. population effect size). The method of effect size calculation is an important contextual factor that must be considered when designing and interpreting research. This has particular importance for interpreting treatment effects and when performing sample size calculations. The observed influence of calculation method also highlights the difficulty of using an effect size in isolation when making a judgement about a measure's general responsiveness.[81]

Important Change

A key finding was that important change values obtained through anchor-based methods (eg, based on patient opinion or comparative measures) were higher than those for the distribution methods (eg, statistical estimates) among the studies that used both approaches.[65–67] In fact, the MCID of one measure (SIS) tripled in magnitude from the distribution to anchor-based approach using the same subjects.[66] Our results indicate that distribution methods result in smaller MCID and researchers may be tempted to interpret their findings in light of this MCID, especially if the intervention has small effects. However, statistically-driven distribution methods have been criticized for their lack of meaning to participants[1,82] The MCID from anchor-based approaches are larger, and may provide a challenge in finding therapies that can achieve this effect. Some have questioned whether it is realistic to achieve MCIDs derived from patient-perceived global rating scales (anchor-based) because many factors (e.g., recall bias, baseline characteristics, expectations of treatment and question format) can affect patients' perception of change.[64,83] It has been suggested that a combination of anchor–based methods from patient and clinical perspectives be used to determine a MCID value and distribution-based methods should only supplement this information.[82]

Minimal Detectable Change

This synthesis highlighted important differences between lab-based and participant - perceived functional measures. Measures that capture perceived function in one's own environment (eg. SIS, MAL) required larger values to surpass their measurement error than

lab-based performance measures (eg. Wolf, ARAT). Although incorporating patient perspectives of functional change is an important component of capturing meaningful outcomes in neurorehabilitation research,[84] researchers should be aware however of the larger sample size required to be able to capture true change using perceived function measures. Quality criteria guidelines recommend that the values needed to surpass measurement error calculated at a 95% confidence interval (ie. true change captured by MDC_{95}) should be less than the minimum values considered to be important (ie. MCID).[85] In contrast, our study found that the MDC_{95} and MDC_{90} for patient-perceived functional measures (eg. MAL, SIS) were similar or greater than the MCID. There remains debate in the literature however concerning the best estimate of minimal detectable change values for self-report measures. Some researchers argue that MDC_{90} and MDC_{95} produce overly conservative estimates and recommend setting the minimal detectable change to one SEM.[86,87]

One of the limitations of this paper is there is no one standard approach for conceptualizing responsiveness. We utilized the broad taxonomy described by Beaton et al.[1] One international group (COSMIN) defined a narrower conceptualization of responsiveness which used a longitudinal validity approach such as correlating change with an external criterion or determining the area under the receiver operating characteristic (ROC) curve to distinguish between known groups.[81] In addition, while there are established appraisal guidelines for intervention studies (e.g., PeDRO Score), consensus has yet to be reached with regards to assessing the rigour of studies that measure an outcome measure's psychometric properties. Application of the Beaton taxonomy, which focuses on the nature of the change, provided a useful framework for understanding the state of the stroke literature in regards to the broader topic of measuring functional change. Effect sizes provided in this study can inform hypotheses for future responsiveness testing. Finally, ceiling and floor effects were also not considered in this review. Large ceiling and/or floor effects may indicate a subsection of the population for which measures are not as responsive and is another factor that can bias estimates of important change values.[88] Given the contextual nature of responsiveness, the dominance of CIMT among the included studies should be noted. Samples from CIMT trials made up 80% of the important change articles, 56% of the detectable change articles and 50% of the RCTs used to examine observed change in response to an effective treatment. Thus, the literature informing the ability of measurement tools to capture functional change in the upper extremity post stroke may be biased to those individuals who are eligible for CIMT trials.

CONCLUSIONS

In summary, this synthesis revealed important findings that have implications for the measurement and interpretation of upper extremity functional recovery following stroke. The magnitude of important change or change that surpasses measurement error varied substantially depending on the method of calculation used. Our findings suggest that rehabilitation treatments can affect patient perceptions of functional change as effectively as lab-based functional measures; however research studies may require higher sample sizes to account for the larger measurement error associated with patient-perceived functional measures. Future studies examining meaningful change in upper extremity function in varied

subgroups of individuals (ie. at different levels of stroke severity and stage of recovery) are needed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by fundings from the Canadian Health Institute of Research (CIHR: 137053) (LAS), the Canadian Health Institute of Research (CIHR MSH-63617) (JJE), and Michael Smith Foundation of Health Research (JJE).

References

1. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. J Clin Epidemiol. 2001; 54:1204–1217. [PubMed: 11750189]
2. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010; 63:737–745. [PubMed: 20494804]
3. Kain ZN, MacLaren J. P less than. 05: What does it really mean? Pediatrics. 2007; 119:608. [PubMed: 17332213]
4. Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database Syst Rev. 2007; (4)
5. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychol Methods. 2002; 7:105–125. [PubMed: 11928886]
6. Stratford PW. Getting more from the literature: estimating the standard error of measurement from reliability studies. Physiother Can. 2004; 56:27–30.
7. Flansbjer UB, Holmbäck AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med. 2005; 37:75–82. [PubMed: 15788341]
8. Angst F. The new COSMIN guidelines confront traditional concepts of responsiveness. BMC Med Res Methodol. 2011; 11:152. [PubMed: 22099330]
9. Hozo S, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005; 5:13. [PubMed: 15840177]
10. Beebe JA, Lang CE. Relationships and responsiveness of six upper extremity function tests during the first six months of recovery after stroke. J Neurol Phys Ther. 2009; 33:96–103. [PubMed: 19556918]
11. Hsueh I, Hsieh C. Responsiveness of two upper extremity function instruments for stroke inpatients receiving rehabilitation. Clin Rehabil. 2002; 16:617–624. [PubMed: 12392336]
12. Lin JH, Hsu MJ, Sheu CF, et al. Psychometric comparisons of 4 measures for assessing upper-extremity function in people with stroke. Phys Ther. 2009; 89:840–850. [PubMed: 19556333]
13. Rabadi MH, Rabadi FM. Comparison of the action research arm test and the fugl-meyer assessment as measures of upper-extremity motor weakness after stroke. Arch Phys Med Rehabil. 2006; 87:962–966. [PubMed: 16813784]
14. Barreca SR, Stratford PW, Masters LM, Lambert CL, Griffiths J, McBay C. Validation of three shortened versions of the Chedoke Arm and Hand Activity Inventory. Physiotherapy Canada. 2006; 58:148–156.
15. Barreca SR, Stratford PW, Lambert CL, Masters LM, Streiner DL. Test-retest reliability, validity, and sensitivity of the Chedoke Arm and Hand Activity Inventory: a new measure of upper-limb function for survivors of stroke. Arch Phys Med Rehabil. 2005; 86:1616–1622. [PubMed: 16084816]

16. Barreca SR, Stratford PW, Masters LM, Lambert CL, Griffiths J. Comparing 2 versions of the Chedoke Arm and Hand Activity Inventory with the Action Research Arm Test. Phys Ther. 2006; 86:245–253. [PubMed: 16445338]
17. Pandyan AD, Cameron M, Powell J, Stott DJ, Granat MH. Contractures in the post-stroke wrist: a pilot study of its time course of development and its association with upper limb recovery. Clin Rehabil. 2003; 17:88–95. [PubMed: 12617383]
18. Brunner IC, Skouen JS, Strand LI. Recovery of upper extremity motor function post stroke with regard to eligibility for constraint-induced movement therapy. Top Stroke Rehabil. 2011; 18:248–257. [PubMed: 21642062]
19. Au-Yeung SS, Hui-Chan CW. Predicting recovery of dextrous hand function in acute stroke. Disabil Rehabil. 2009; 31:394–401. [PubMed: 18608431]
20. Rand D, Eng JJ. Disparity between functional recovery and daily use of the upper and lower extremities during subacute stroke rehabilitation. Neurorehabil Neural Repair. 2012; 26:76–84. [PubMed: 21693771]
21. Roiha K, Kirveskari E, Kaste M, et al. Reorganization of the primary somatosensory cortex during stroke recovery. Clin Neurophysiol. 2011; 122:339–345. [PubMed: 20673646]
22. Blennerhassett JM, Avery RM, Carey LM. The test-retest reliability and responsiveness to change for the Hand Function Survey during stroke rehabilitation. Aust Occup Ther J. 2010; 57:431–438. [PubMed: 21091710]
23. Rehme AK, Fink GR, von Cramon DY, Grefkes C. The role of the contralesional motor cortex for motor recovery in the early days after stroke assessed with longitudinal FMRI. Cereb Cortex. 2011; 21:756–768. [PubMed: 20801897]
24. Filiatrault J, Arsenault AB, Dutil E, Bourbonnais D. Motor function and activities of daily living assessments: a study of three tests for persons with hemiplegia. Am J Occup Ther. 1991; 45:806–810. [PubMed: 1928288]
25. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The Stroke Impact Scale version 2.0. evaluation of reliability, validity, and sensitivity to change. Stroke. 1999; 30:2131–2140. [PubMed: 10512918]
26. Kamel A, Ghani AA, Zaiton MA, El-Motayam AS, El-Fattah DA. Health related quality of life in stroke survivors measured by the Stroke Impact Scale. Egypt J of Neurol Psychiat Neurosurg. 2010; 47:267–274.
27. Sezer N, Yavuzer G, Sivrioglu K, Basaran P, Koseoglu BF. Clinimetric properties of the Duruoz Hand Index in patients with stroke. Arch Phys Med Rehabil. 2007; 88:309–314. [PubMed: 17321822]
28. Bouffioulx É, Arnould C, Thonnard JL. Satisfaction with activity and participation and its relationships with body functions, activities, or environmental factors in stroke patients. Arch Phys Med Rehabil. 2011; 92s:1404–1410.
29. Rand D, Gottlieb D, Weiss P. Recovery of patients with a combined motor and proprioception deficit during the first six weeks of post stroke rehabilitation. Phys Occup Ther Geratr. 2001; 18:69–87.
30. Wittenberg GF, Bastings EP, Fowlkes AM, Morgan TM, Good DC, Pons TP. Dynamic course of intracortical TMS paired-pulse responses during recovery of motor function after stroke. Neurorehabil Neural Repair. 2007; 21:568–573. [PubMed: 17522261]
31. Higgins J, Mayo NE, Desrosiers J, Salbach NM, Ahmed S. Upper-limb function and recovery in the acute phase poststroke. J Rehabil Res Dev. 2005; 42:65–76.
32. Mayo NE, Wood-Dauphinee S, Ahmed S, et al. Disablement following stroke. Disabil Rehabil. 1999; 21:258–268. [PubMed: 10381238]
33. Desrosiers J, Malouin F, Richards C, Bourbonnais D, Rochette A, Bravo G. Comparison of changes in upper and lower extremity impairments and disabilities after stroke. Int J Rehabil Res. 2003; 26:109–116. [PubMed: 12799604]
34. Feydy A, Carlier R, Roby-Brami A, et al. Longitudinal study of motor recovery after stroke: Recruitment and focusing of brain activation. Stroke. 2002; 33:1610–1617. [PubMed: 12053000]

35. Hsu SS, Hu MH, Wang YH, Yip PK, Chiu JW, Hsieh CL. Dose-response relation between neuromuscular electrical stimulation and upper-extremity function in patients with stroke. Stroke. 2010; 41:821–824. [PubMed: 20203321]

36. Page SJ, Levine P, Leonard AC. Modified constraint-induced therapy in acute stroke: A randomized controlled pilot study. Neurorehabil Neural Repair. 2005; 19:27–32. [PubMed: 15673841]

37. Sun SF, Hsu CW, Sun HP, Hwang CW, Yang CL, Wang JL. Combined botulinum toxin type A with modified constraint-induced movement therapy for chronic stroke patients with upper extremity spasticity: a randomized controlled study. Neurorehabil Neural Repair. 2010; 24:34–41. [PubMed: 19729582]

38. Myint JM, Yuen GF, Yu TK, et al. A study of constraint-induced movement therapy in subacute stroke patients in Hong Kong. Clin Rehabil. 2008; 22:112–124. [PubMed: 18212033]

39. van der Lee JH, Wagenaar RC, Lankhorst GJ, Vogelaar TW, Deville WL, Bouter LM. Forced use of the upper extremity in chronic stroke patients: results from a single-blind randomized clinical trial. Stroke. 1999; 30:2369–2375. [PubMed: 10548673]

40. Harris JE, Eng JJ, Miller WC, Dawson AS. A self-administered graded repetitive arm supplementary program (GRASP) improves arm function during inpatient stroke rehabilitation: A multi-site randomized controlled trial. Stroke. 2009; 40:2123–2128. [PubMed: 19359633]

41. Shindo K, Fujiwara T, Hara J, et al. Effectiveness of hybrid assistive neuromuscular dynamic stimulation therapy in patients with subacute stroke. Neurorehabil Neural Repair. 2011; 25:830–837. [PubMed: 21666139]

42. Dahl AE, Askim T, Stock R, Langorgen E, Lydersen S, Indredavik B. Short- and long-term outcome of constraint-induced movement therapy after stroke: a randomized controlled feasibility trial. Clin Rehabil. 2008; 22:436–447. [PubMed: 18441040]

43. Cacchio A, De Blasis E, De Blasis V, Santilli V, Spacca G. Mirror therapy in complex regional pain syndrome type 1 of the upper limb in stroke patients. Neurorehabil Neural Repair. 2009; 23:792–799. [PubMed: 19465507]

44. Kowalczewski J, Gritsenko V, Ashworth N, Ellaway P, Prochazka A. Upper-extremity functional electric stimulation-assisted exercises on a workstation in the subacute phase of stroke recovery. Arch Phys Med Rehabil. 2007; 88:833–839. [PubMed: 17601461]

45. 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. [PubMed: 16401430]

46. Taub E, Uswatte G, King DK, Morris D, Crago JE, Chatterjee A. A placebo-controlled trial of constraint-induced movement therapy for upper extremity after stroke. Stroke. 2006; 37:1045–1049. [PubMed: 16514097]

47. Wittenberg GF, Chen R, Ishii K, et al. Constraint-induced therapy in stroke: magnetic-stimulation motor maps and cerebral activation. Neurorehabil Neural Repair. 2003; 17:48–57. [PubMed: 12645445]

48. Wu CY, Chuang LL, Lin KC, Chen HC, Tsay PK. 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; 25:130–139. [PubMed: 20947493]

49. Gauthier LV, Taub E, Perkins C, Ortmann M, Mark VW, Uswatte G. Remodeling the brain: plastic structural brain changes produced by different motor therapies after stroke. Stroke. 2008; 39:1520–1525. [PubMed: 18323492]

50. Khan CM, Oesch PR, Gamper UN, Kool JP, Beer S. Potential effectiveness of three different treatment approaches to improve minimal to moderate arm and hand function after stroke: a pilot randomized clinical trial. Clin Rehabil. 2011; 25:1032–1041. [PubMed: 21788267]

51. Tariah HA, Almalty A, Sbeih Z, Al-Oraibi S, Bernhardt J, Rowe V. Constraint induced movement therapy for stroke survivors in Jordon: a home-based model. Int J Ther Rehabil. 2010; 17:638–646.

52. Dromerick AW, Lang CE, Birkenmeier RL, et al. Very early constraint-induced movement during stroke rehabilitation (VECTORS): A single-center RCT. Neurology. 2009; 73:195–201. [PubMed: 19458319]

53. Church C, Price C, Pandyan AD, Huntley S, Curless R, Rodgers H. Randomized controlled trial to evaluate the effect of surface neuromuscular electrical stimulation to the shoulder after acute stroke. Stroke. 2006; 37:2995–3001. [PubMed: 17053181]
54. Wu CY, Chen CL, Tsai WC, Lin KC, Chou SH. A randomized controlled trial of modified constraint-induced movement therapy for elderly stroke survivors: changes in motor impairment, daily functioning, and quality of life. Arch Phys Med Rehabil. 2007; 88:273–278. [PubMed: 17321816]
55. Kimberley TJ, Lewis SM, Auerbach EJ, Dorsey LL, Lojovich JM, Carey JR. Electrical stimulation driving functional improvements and cortical changes in subjects with stroke. Exp Brain Res. 2004; 154:450–460. [PubMed: 14618287]
56. Lin KC, Chang YF, Wu CY, Chen YA. Effects of constraint-induced therapy versus bilateral arm training on motor performance, daily functions, and quality of life in stroke survivors. Neurorehabil Neural Repair. 2009; 23:441–448. [PubMed: 19118130]
57. Lin KC, Wu CY, Liu JS, Chen YT, Hsu CJ. Constraint-induced therapy versus dose-matched control intervention to improve motor ability, basic/extended daily functions, and quality of life in stroke. Neurorehabil Neural Repair. 2009; 23:160–165. [PubMed: 18981188]
58. Ertelt D, Small S, Solodkin A, et al. Action observation has a positive impact on rehabilitation of motor deficits after stroke. Neuroimages. 2007; 36:T164–T173.
59. Conroy SS, Whitall J, Dipietro L, et al. Effect of gravity on robot-assisted motor training after chronic stroke: a randomized trial. Arch Phys Med Rehabil. 2011; 92:1754–1761. [PubMed: 21849168]
60. Liao WW, Wu CY, Hsieh YW, Lin KC, Chang WY. Effects of robot-assisted upper limb rehabilitation on daily function and real-world arm activity in patients with chronic stroke: a randomized controlled trial. Clin Rehabil. 2012; 26:111–120. [PubMed: 21840917]
61. Page SJ, Levin L, Hermann V, Dunning K, Levine P. Longer versus shorter daily durations of electrical stimulation during task-specific practice in moderately impaired stroke. Arch Phys Med Rehabil. 2012; 93:200–206. [PubMed: 22289227]
62. Hsieh YW, Wu CY, Liao WW, Lin KC, Wu KY, Lee CY. Effects of treatment intensity in upper limb robot-assisted therapy for chronic stroke: A pilot randomized controlled trial. Neurorehabil Neural Repair. 2011; 25:503–511. [PubMed: 21436390]
63. Fritz SL, George SZ, Wolf SL, Light KE. Participant perception of recovery as criterion to establish importance of improvement for constraint-induced movement therapy outcome measures: a preliminary study. Phys Ther. 2007; 87:170–178. [PubMed: 17244694]
64. Lang CE, Edwards DF, Birkenmeier RL, Dromerick AW. Estimating minimal clinically important differences of upper-extremity measures early after stroke. Arch Phys Med Rehabil. 2008; 89:1693–1700. [PubMed: 18760153]
65. Wang TN, Lin KC, Wu CY, Chung CY, Pei YC, Teng YK. Validity, responsiveness, and clinically important difference of the ABILHAND questionnaire in patients with stroke. Arch Phys Med Rehabil. 2011; 92:1086–1091. [PubMed: 21704789]
66. Lin KC, Fu T, Wu CY, et al. Minimal detectable change and clinically important difference of the Stroke Impact Scale in stroke patients. Neurorehabil Neural Repair. 2010; 24:486–492. [PubMed: 20053950]
67. Lin KC, Hsieh YW, Wu CY, Chen CL, Jang Y, Liu JS. Minimal detectable change and clinically important difference of the Wolf Motor Function Test in stroke patients. Neurorehabil Neural Repair. 2009; 23:429–434. [PubMed: 19289487]
68. Fritz SL, Blanton S, Uswatte G, Taub E, Wolf SL. Minimal detectable change scores for the Wolf Motor Function Test. Neurorehabil Neural Repair. 2009; 23:662–667. [PubMed: 19498013]
69. Wolf SL, Thompson PA, Morris DM, et al. The EXCITE trial: attributes of the Wolf Motor Function Test in patients with subacute stroke. Neurorehabil Neural Repair. 2005; 19:194–205. [PubMed: 16093410]
70. Uswatte G, Giuliani C, Winstein C, Zeringue A, Hobbs L, Wolf SL. Validity of accelerometry for monitoring real-world arm activity in patients with subacute stroke: evidence from the extremity constraint-induced therapy evaluation trial. Arch Phys Med Rehabil. 2006; 87:1340–1345. [PubMed: 17023243]

71. Uswatte G, Taub E, Morris D, Vignolo M, McCulloch K. Reliability and validity of the upper-extremity Motor Activity Log-14 for measuring real-world arm use. Stroke. 2005; 36:2493–2496. [PubMed: 16224078]
72. 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; 67:1189–1194. [PubMed: 17030751]
73. Kopp B, Kunkel A, Flor H, et al. The Arm Motor Ability Test: reliability, validity, and sensitivity to change of an instrument for assessing disabilities in activities of daily living. Arch Phys Med Rehabil. 1997; 78:615–620. [PubMed: 9196469]
74. Richards L, Stoker-Yates J, Pohl P, Wallace D, Duncan P. Reliability and validity of two tests of upper extremity motor function post-stroke. Occup Ther J Res. 2001; 21:201–219.
75. Carod-Artal FJ, Coral LF, Trizotto DS, Moreira CM. The Stroke Impact Scale 3. 0: Evaluation of acceptability, reliability, and validity of the Brazilian version. Stroke. 2008; 39:2477–2484. [PubMed: 18635846]
76. van der Lee JH, Beckerman H, Knol DL, De Vet HC, Bouter LM. Clinimetric properties of the Motor Activity Log for the assessment of arm use in hemiparetic patients. Stroke. 2004; 35:1410. [PubMed: 15087552]
77. van der Lee JH, Beckerman H, Lankhorst GJ, Bouter LM. The responsiveness of the Action Research Arm Test and the Fugl-Meyer assessment scale in chronic stroke patients. J Rehabil Med. 2001; 33:110–113. [PubMed: 11482350]
78. Lydick E, Epstein RS. Interpretation of quality of life changes. Qual Life Res. 1993; 2:221–226. [PubMed: 8401458]
79. Biernaskie J, Chernenko G, Corbett D. Efficacy of rehabilitative experience declines with time after focal ischemic brain injury. J Neurosci. 2004; 24:1245–1254. [PubMed: 14762143]
80. Dobkin BH, Dorsch A. The promise of mHealth: daily activity monitoring and outcome assessments by wearable sensors. Neurorehabil Neural Repair. 2011; 25:788–798. [PubMed: 21989632]
81. Mokkink L, Terwee C, Knol D, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010; 10:22. [PubMed: 20298572]
82. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008; 61:102–109. [PubMed: 18177782]
83. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. J Clin Epidemiol. 2002; 55:900–908. [PubMed: 12393078]
84. Salter K, Jutai J, Teasell R, Foley N, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF activity. Disabil Rehabil. 2005; 27:315–340. [PubMed: 16040533]
85. Terwee CB, Bot SDM, De Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007; 60:34–42. [PubMed: 17161752]
86. Wyrwich KW, Tierney WM, Wolinksy FD. Further evidence supporting standard error of measurement based criterion for identifying meaningful intra-individual change in health-related quality of life. J Clin Epidemiol. 1999; 52:861–873. [PubMed: 10529027]
87. Turner D, Schunemann HJ, Griffith LE, et al. The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol. 2010:63-28-36.
88. Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH. Interpretation of change scores in ordinal clinical scales and health status measures: The whole may not equal the sum of the parts. J Clin Epidemiol. 1996; 49:711–717. [PubMed: 8691219]
89. Streiner, DLNG. Health measurement scales: A practical guide to their development and use. 4. New York: Oxford University Press; 2008.

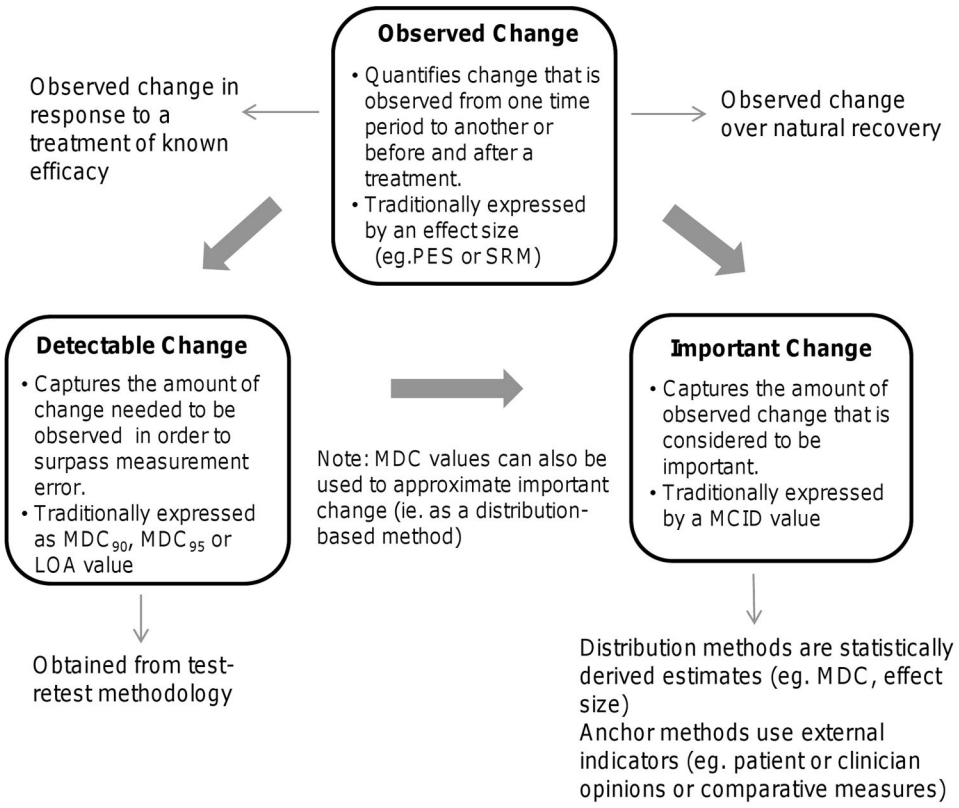

Figure 1.

Graphic representation of observed, important and detectable change

Abbreviations: MDC90: minimal detectable change (with 90% confidence interval); MDC95: minimal detectable change (with 95% confidence interval); LOA: limits of agreement; PES: population effect size; SRM: standardized response mean; MCID: minimally clinically important difference

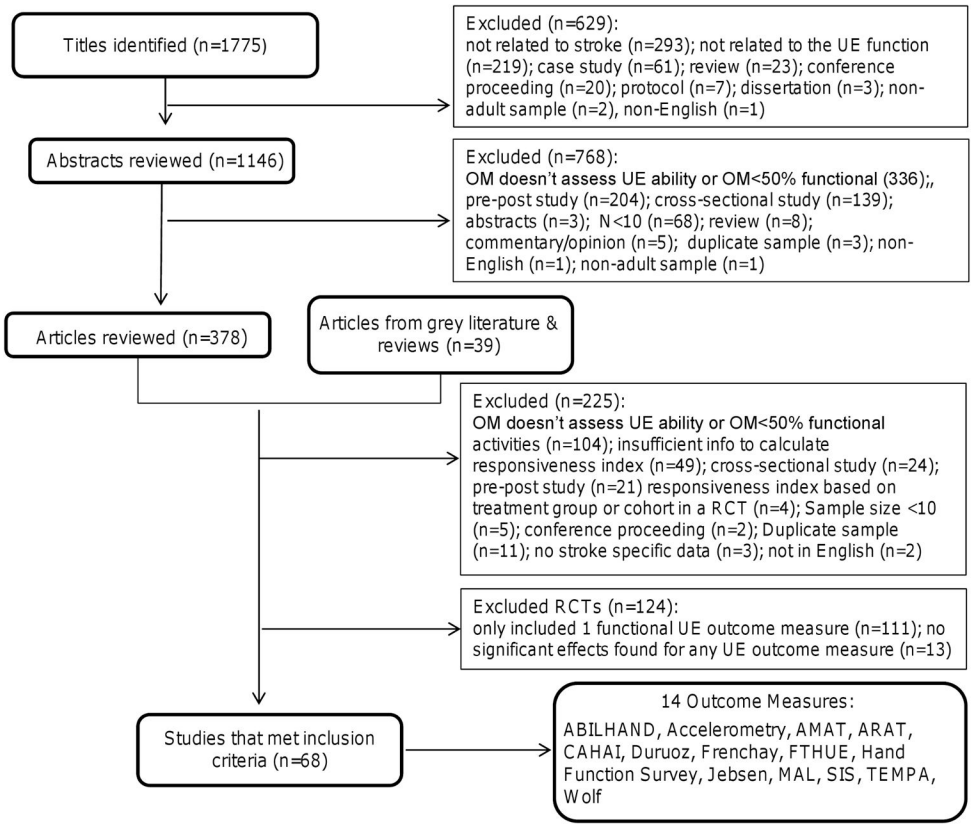

Figure 2.

Flow diagram of process to select final list of outcome measures

Abbreviations: OM: outcome measure; AMAT: Arm Motor Activity Test; ARAT: Action Research Arm Test; CAHAI: Chedoke Arm and Hand Activity Inventory; Duruoz: Durouz Hand Index; Frenchay: Frenchay Arm Test; FTHUE: Functional Test for the Hemiplegic Upper Extremity; Jebsen: Jebsen Hand Function Test; MAL: Motor Activity Log; SIS: Stroke Impact Scale; TEMPA: Upper Extremity Performance Scale for the Elderly; Wolf: Wolf Motor Function Test

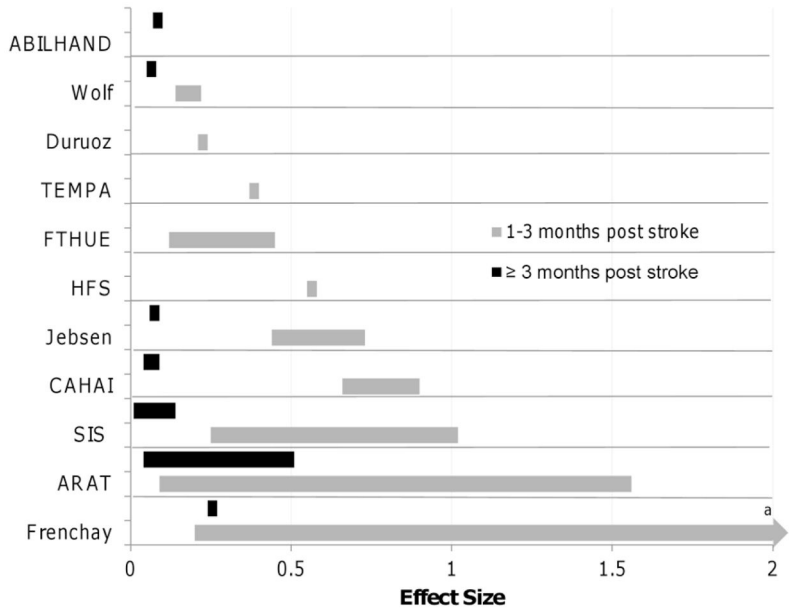

Figure 3.

Effect sizes by measure calculated at <3 months and ≥ 3 months post stroke

The bars on the graph represent the range of effect sizes calculated from studies that measured UE function across time. [a]the full range of the effect sizes for the Frenchay is 0.2–5.

Abbreviations: Wolf: Wolf Motor Function Test; Duruoz: Duruoz Hand Index; TEMPA: Upper Extremity Performance Scale for the Elderly; FTHUE: Functional Test for the Hemiplegic Upper Extremity; Jebsen: Jebsen Hand Function Test; CAHAI: Chedoke Arm and Hand Inventory; SIS: Stroke Impact Scale (hand scale); ARAT: Action Research Arm Test; Frenchay: Frenchay Arm Test

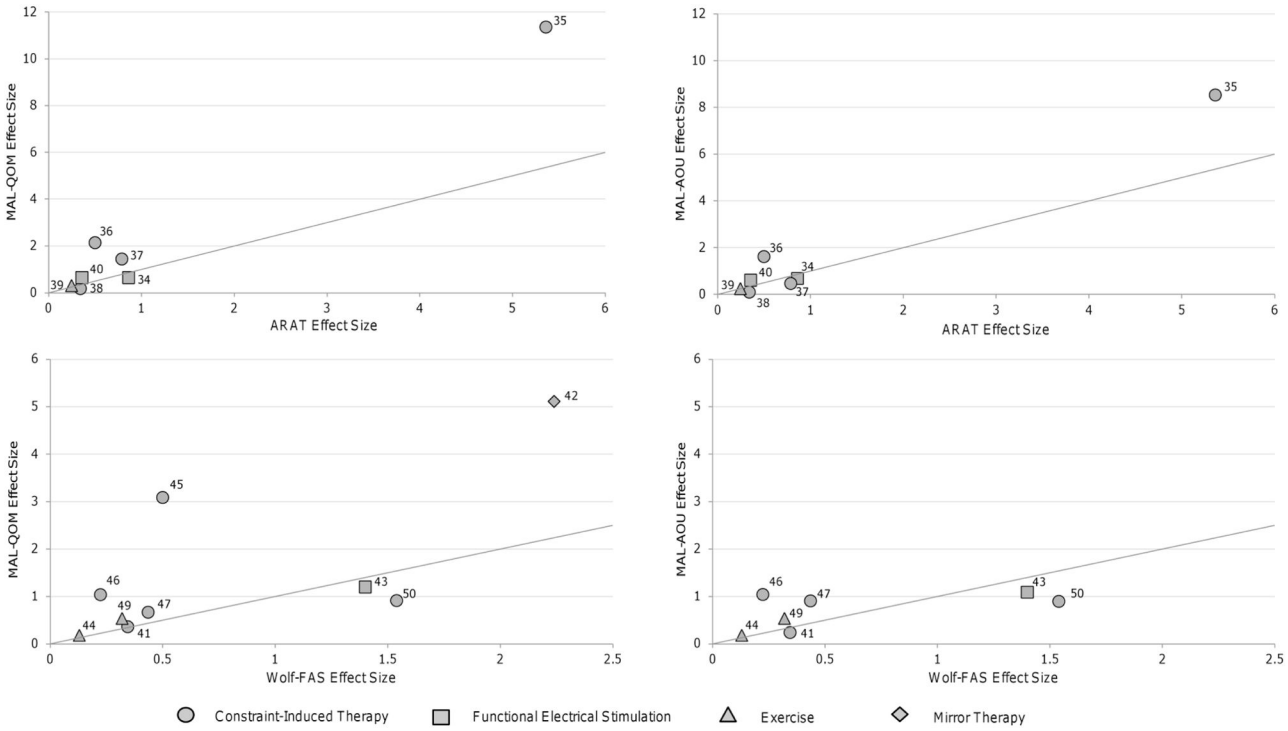

Figure 4.

Comparison of observed change captured by 'lab-based' versus 'patient-perceived functional measures

Points on the graph represent the effect sizes obtained from a single study. Lines on the graph represent a 1:1 relationship between the 'lab-based' vs 'patient-perceived' functional measures. Lab-based measures are located on the X-axes (ie. Wolf, ARAT). Patient-perceived functional measures are located on the Y axes (ie. MAL).

Abbreviations: MAL: Motor Activity Log; AOU: amount of use scale; QOM: quality of movement scale; CIMT: ARAT: Action Research Arm Test; Wolf: Wolf Motor Function Test; FAS: functional ability scale

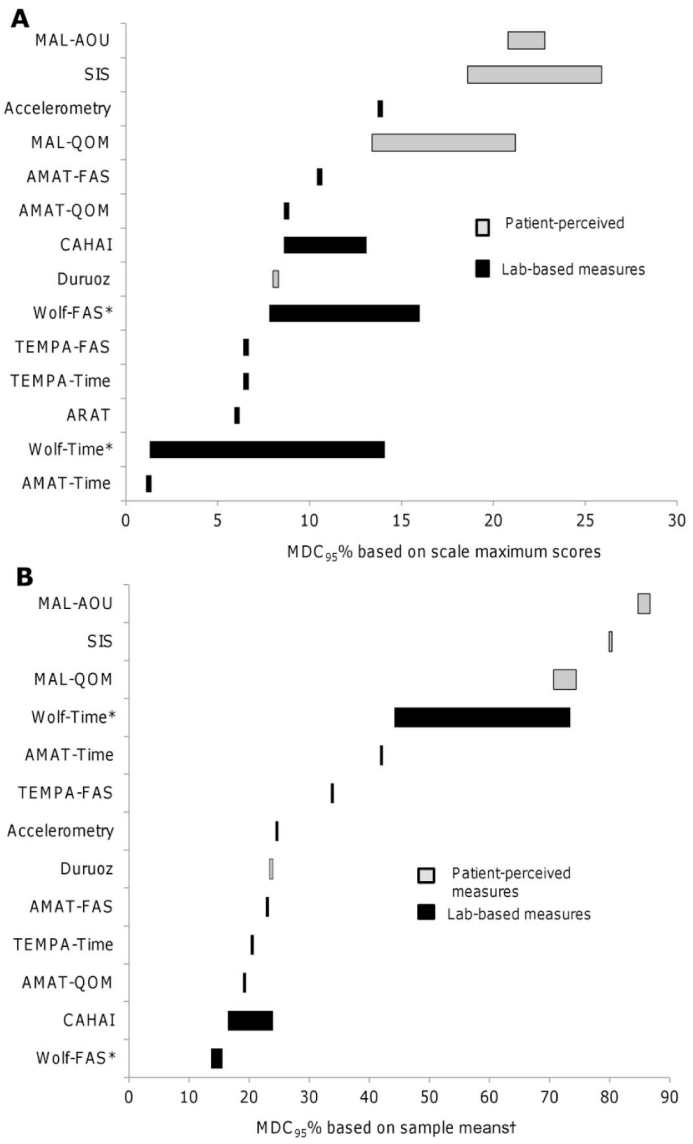

Figure 5.

Comparison of detectable change (calculated at 95% confidence level) relative to the sample means and scale maximums

The bars beside each measure represent the range of MDC_{95}%values extracted or calculated from different studies. *Two studies analysed different subsets of the same sample to obtain multiple estimates. †Estimates for the ARAT, Wolf (FAS) and SIS are missing from this graph as sample means were not provided in two studies.[11,65]

Abbreviations: MDC_{95}%: Minimally detectable change (with a 95% confidence interval) expressed as a percentage of A: the scale maximum score and B: the sample means; MAL: Motor Activity Log; AOU: amount of use scale; SIS: Stroke Impact Scale (hand); QOM: quality of movement scale; TEMPA: Upper Extremity Performance Scale for the Elderly; AMAT: Arm Motor Ability Test; FAS: functional

Table 1

Minimally clinically important difference (MCID) values calculated by distribution and anchor-based methods[a]

	ABILHAND	Acceler	ARAT	MAL QOM	MAL AOU	SIS	Wolf FAS	Wolf Time
MCID obtained by distribution-based methods[b]								
Effect size	0.26[65] (in logits)					5.8[66] (5.8%)	0.14[67] (2.8%)	1.4[67] (1.1%)
MICD obtained by anchor-based methods[c]								
Fugl-Meyer							0.33[67] (6.6%)	1.64[67] (1.4%)
% Recovery	0.35[65] (in logits)				NS[63]	17.8[66] (17.8%)		NS[63]
Global		NS[64]	12–17[64] (21–30%)	1.0–1.1[64] (20–22%)			1.0–1.2[64] (20–24%)	19.0[64] (16%)

[a] MCID values are displayed in raw scores (% of scale maximum score)

[b] An effect size of 0.2 was used as the distribution method to determine MCID

[c] Anchor-based methods used to calculate MCID consisted of: Fugl-Meyer, change of 6–10 points on the Fugl-Meyer upper extremity scale; % Recovery, 10–15% or 50% recovery on the Stroke Impact Scale global recovery item; Global, perception of important change on a global rating scale

Abbreviations: Accelerom: Accelerometry; ARAT: Action Research Arm Test; MAL: Motor Activity Log; QOM: quality of movement scale; AOU: amount of use scale; SIS: Stroke Impact Scale; Wolf: Wolf Motor Function Test; FAS: functional ability scale; NS: MCID values could not be calculated due to no relationship between change scores and global recovery scales; logits: log odds scale units that allow Likert scale scores (i.e. ABILHAND raw scores) to be interpreted as interval scores.[89]