A META-ANALYSIS OF CONSTRAINT-INDUCED MOVEMENT THERAPY AFTER STROKE

Gyrd Thrane, Cand. San.1,2, Oddgeir Friborg, PhD3, Audny Anke, MD, PhD2,4 and Bent Indredavik, MD, PhD5,6

From the 1Department of Health and Care Sciences, University of Tromsø, 2Department of Rehabilitation, University Hospital of North Norway, 3Department of Psychology, 4Faculty of Health Sciences, Department of Clinical Medicine, University of Tromsø, Tromsø, 5Department of Neuroscience, Faculty of Medicine, Norwegian University of Science and Technology and 6Stroke Unit, St Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

Objective: To evaluate the effect of constraint-induced movement therapy in adult stroke patients and to examine the impact of time since stroke and various treatment modalities.

Data sources: PubMed, EMBASE, Cochrane and PEDro trial registers were searched for clinical trials published before November 2012.

Study selection: Randomized or quasi-randomized controlled trials of constraint-induced movement therapy lasting 2–7 h/day for 8–28 days were included.

Data extraction: Measurements were classified into the following categories: arm motor function, arm motor activity, activities of daily living, and participation. A pooled standardized mean difference (SMD) was calculated for each category. Moderators were: trial quality, behavioural techniques, amount of training, time since stroke, shaping, and the nature of the control group.

Data synthesis: Of 3,842 records initially screened 23 trials were included. A small post-treatment effect was found on arm motor function (SMD 0.28, 95% confidence interval (CI) 0.11–0.44). Meanwhile, a moderate effect on arm motor activity was found post-treatment (SMD 0.51, 95% CI 0.30–0.73) and at 3–6 months follow-up (SMD 0.41, 95% CI 0.08–0.74).

Conclusion: Constraint-induced movement therapy can improve arm motor function and improve arm motor activities and may have a lasting effect on arm motor activity.

Key words: constraint-induced movement therapy; adult; cerebral stroke; meta-analysis; ICF.


INTRODUCTION

Paresis of the upper extremity is a common impairment after stroke. Even after recovery approximately 20% of survivors cannot independently complete upper extremity activities, such as feeding and personal care (1). Constraint-induced movement therapy (CIMT) is a treatment for mild and moderate upper extremity motor dysfunction after stroke. The treatment consists of intensive exercises for the more affected upper limb. In addition, different behavioural procedures, such as a behavioural contract, systematic feedback and encouragement of real-world problem-solving, can be used to enhance the transfer of gained motor skills into daily activities. The less affected arm is also constrained by a mitt or cast during the course of treatment (2, 3). The original protocol included 10 days of therapy for 6 h a day and was developed for chronic stroke patients (2).

One large clinical trial (4) and 5 systematic reviews on the effect of CIMT have been published (5–9). In the large clinical trial, the effect of traditional CIMT 3–9 months post-stroke was investigated in 222 chronic stroke patients from 7 American sites (4). Significant effects on arm motor function, self-reported arm use, and self-reported quality of arm use were found. A Cochrane systematic review with a subsequent update found reduced arm motor impairment (standardized mean difference (SMD) 0.65, 95% confidence interval (CI) 0.15–1.15), together with increased arm motor function (SMD 0.44, 95% CI 0.03–0.84) and perceived arm motor function (SMD 1.16, 95% CI 1.05–1.27) after CIMT (7, 10). Two other recent reviews found an effect of modified CIMT compared with traditional rehabilitation on the Fugl-Meyer assessment (mean difference (MD)=7.80, 95% CI 4.21–11.38), Action Research Arm Test (MD=14.15, 95% CI 10.71–17.59), Wolf Motor Function Test (MD=0.46, 95% CI 0.33–0.59), Functional Independence Measure (MD=7.00, 95% CI 0.75–13.26) and Motor Activity Log (MAL) (MD=0.85, 95% CI 0.62–1.08) (8, 9).

A significant variety of CIMT modifications is present in the various trials, which is also evident in the heterogeneity shown in earlier reviews (6, 7, 9). CIMT is a multifaceted treatment that combines functional training and behavioural treatment. In particular, the behavioural procedures are used in different ways across trials. Moreover, the trials differ in terms of the nature of the control group, the time in therapy, and the time from stroke to inclusion. All earlier systematic reviews state that the empirical evidence is limited and underpowered. There is only one meta-analysis of follow-up data, which included only 2 trials (7). Even though the summary effect of CIMT on single measurements is well studied in several meta-analyses (5, 8, 9), the effect on groups of measurements in the Inter-
national Classification of Functioning, Disability and Health (ICF) dimensions of body function, activity, and participation have only been partially examined (7). None of the analyses have combined more than one measurement from each study in their analysis. The ideal time to start the treatment after stroke is also uncertain. More information is therefore needed on the effect of CIMT within each ICF domain and the effect of follow-up and various factors on the therapy outcome.

The aims of this systematic review were to combine the measurements used in CIMT trials in a valid manner according to the ICF constructs in order to: (i) evaluate the effect of CIMT on arm motor function, arm motor activities, activities of daily living (ADL), and participation immediately after treatment and at 3–6 months follow-up; and (ii) assess how the reported effects are associated with the quality of the trial, the use of behavioural techniques, hours of training, the nature of the control group, and the time from stroke to trial enrolment.

METHODS

Data sources

PubMed, EMBASE, and CINAHL databases, as well as the Cochrane and PEDro Trial registers, were searched for empirical papers published from the inception of each database up to November 2012. The key words used for identifying CIMT trials were: CIMT, mCIMT, Forced use, constraint-induced movement therapy. The term “Forced use” was included because this was the original term used in the preliminary CIMT experiments (2). From the PEDro trial register all records containing 1 of these keywords were reviewed. In the other database searches the CIMT search were combined with the following search terms for identifying stroke trials: stroke, hemiplegia, hemiparesis, hemiplegic, cerebrovascular accident, cerebrovascular accident (CVA), apoplexy, apoplectic or vascular accident brain. A research assistant read all the identified titles and excluded obviously irrelevant papers. In addition, the reference list from previous systematic reviews and included papers were assessed for relevant records. Two of the authors (GT and AA) read through the remaining abstracts and marked them as relevant, irrelevant, or potentially relevant based on the inclusion criteria. Full-text articles were retrieved for papers ranked as potentially relevant. Papers were included if: (i) a randomized controlled trial or quasi-randomized controlled trial was performed; (ii) patients were >18 years old and had a diagnosis of ischaemic or haemorrhagic stroke; (iii) CIMT, mCIMT or forced use was compared with a control group; (iv) the treatment dose in the experimental group was 2–7 h of training per treatment day for a period of 8–28 days; and (v) there were more than 4 subjects in the CIMT group. Papers were excluded if: (i) the CIMT group received additional interventions (e.g., botulinum toxin), or (ii) the control group received a different form of CIMT. Disagreements were resolved by a consensus discussion between the 2 assessors.

Data extraction

Data were extracted from each paper to a coding form by a research assistant. Author name, publication year, and country were recorded for each paper. Pre-test, post-test, and 3–6 month follow-up data were extracted to compute the effect sizes. The PEDro score was obtained from the PEDro database (11) as a measure of trial quality. The PEDro Review group assess trials relevant for physiotherapy according and the results are available from the PEDro website (www.pedro.org.au). The PEDro scale is an 11-item assessment of trial quality based on whether the trial reports eligibility criteria, randomization procedure, concealed allocation, blinding of subjects, blinding of therapist, blinding of assessors, adequate follow-up, intention to treat analysis, between-group statistical comparisons and between-group comparability. The items assessing internal validity (2–11) is summed as a score for trial quality, with a possible range of 0–10. The reliability of the total PEDro score is described as “fair” to “good” (12). One trial did not have an official PEDro (13) score and was scored by collaboration between 2 of the authors (GT and AA). If the author stated explicitly that the groups compared in the trial were established by random allocation, then the trial was classified as a randomized controlled trial (RCT). If a trial did not state explicitly that it was randomized, or if the trial used a quasi-random method of allocation, it was classified as controlled clinical trial (CCT) (14). Trials in which the randomization was compromised were also classified as CCTs. According to our pre-specified criteria, we attempted to categorize the time since stroke into 3 categories (<3, 3–12, and >12 months) based on the trial inclusion criteria. However, in practice, this categorization was not possible because most trials included patients from 2 or more of these groups. Therefore, trials with a mean time since stroke of ≤45 days were classified as early intervention, while those with a mean time since stroke >45 days was classified as sub-acute and chronic. The recording of the behavioural techniques (adherence-enhancing techniques and shaping) used in the CIMT was based on the description of the intervention. Shaping was recorded as used if the exercises included motor tasks of short trials successively progressing in task difficulty, and included systematic feedback on performance. The number of adherence-enhancing techniques (behavioural contract, home diary, home practice, motor activity log) was noted. The nature of the control group was categorized into 4 (no treatment, usual care, calibrated intensive or other control treatment). The calibrated intensive category was used when the time in treatment in the control group was equal to that in the CIMT group. Control groups that received an unspecified normal treatment programme were defined as usual care. Other control treatments included non-calibrated experimental treatments. One of the authors transferred the data from the coding form to the computer program. During this process each value was cross-checked with the information from the actual paper. Two authors (GT and AA) classified in consensus the outcome measures according to the International Classification of Functioning, Disability, and Health (ICF) dimensions of body function, activity, and participation. The final classification is shown in Table I. Measurements used to examine the control of voluntary movement (ICF Chapter B760) (15), strength, or muscle tone were assigned to the arm motor function group (body function). The activity outcome dimension was divided into 2 groups. One group contained measures specifically designed to measure arm motor activity (Chapters d430, d440, and d445). The other group contained measures not specific to arm motor activity measuring mobility (Chapter d4), self-care (Chapter d5), and domestic life (Chapter d6) (15). Measures belonging to this group were classified as ADL. As one study reported the use of the MAL as a part of treatment (13); the MAL scores were therefore omitted from the analyses in the present study. If a study reported both the quality of movement and the amount of use scales, only the MAL amount of use scores was used for analysis. Measurements clearly measuring participation were classified in the participation group.

Statistical analyses

The Comprehensive Meta-Analysis (CMA) version 2.2.064 software package was used for statistical analyses. The effect size for each outcome measures was calculated as a standardized mean difference (SMD) score, and was based on the post-test mean differences between the treatment and the control group, adjusted for the pre-test scores. If only post-data were available, the SMD was based on the post-test scores only. One study (16, 17) reported Cohen’s d’s for the outcomes; these values were entered directly into the program. A composite SMD was calculated and the variance of all the outcome measures within each of the 4 measurement domains (Table I) were then calculated in Microsoft Excel according to the formula given by Borenstein (18):

$$\text{var} \left( \frac{1}{m} \sum_{i=1}^{m} Y_i \right) = \left( \frac{1}{m} \right)^2 \text{var} \left( \sum_{i=1}^{m} Y_i \right) = \left( \frac{1}{m} \right)^2 \sum_{i=1}^{m} \text{var} Y_i + \sum_{i<j} r_{ij} \text{cov}(Y_i, Y_j)$$

Where m is the number of measurements Y included in the composite variance, r_{ij} is the correlation between two measurements, Y_i and Y_j, and \text{cov}(Y_i, Y_j) is the respective variances of Y_i and Y_j.
Table I. Measurements used in constraint-induced movement therapy (CIMT) trials and included in the meta-analysis were classified as body function, activity, or participation. Arm motor function, arm motor function, activities of daily living (ADL), and participation were the categories used in the analysis.

<table>
<thead>
<tr>
<th>Body function</th>
<th>Activity</th>
<th>Activities of daily living</th>
<th>Participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action Research Arm Test</td>
<td>Actual Amount of Use Test</td>
<td>Barthel index</td>
<td>Stroke Impact Scale –</td>
</tr>
<tr>
<td>Fugl-Meyer Motor Assessment</td>
<td>Grooved Pegboard Test</td>
<td>Functional Independence Measure</td>
<td>Participation</td>
</tr>
<tr>
<td>Motor Evaluation Scale for Arm in Stroke Patients</td>
<td>Nine-hole Peg Test</td>
<td>Nottingham Extended Activities of Daily Living</td>
<td>Frenchay Activities Index</td>
</tr>
<tr>
<td>Modified Ashford Scale</td>
<td>16HPT</td>
<td>Rehabilitation Activities Profile</td>
<td></td>
</tr>
<tr>
<td>Wolf motor function test – Grip strength ratio</td>
<td>Motor Activity Log – Amount of use</td>
<td>Stroke Impairment Scale – activities</td>
<td></td>
</tr>
<tr>
<td>Wolf motor function test – Strength</td>
<td>Motor Activity Log – Quality of movement</td>
<td>of daily living</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroke Impairment Scale – Hand</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wolf motor function test – Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wolf motor function test – Quality of movement</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As none of the studies reported the correlations between the outcome measures, they were imputed based on the reported correlations from available studies. We used the highest correlation reported between the measurements in a stroke population, which is a conservative choice because high correlations produce wider confidence intervals than lower correlations (18). When correlations between measurements were not found, they were imputed based on correlations between scales of a similar nature. Random effects model meta-analysis was used to calculate the pooled SMDs, the 95% confidence intervals (95% CI), and the p-values within each domain. The combined effects in the 4 studies that had 2 intervention or control-groups (19-22) were calculated using the mean of selected comparisons function of the CMA. This procedure corrects the variance in order to take into account the relationship between the results in the 2 groups. Heterogeneity in effect sizes was examined by calculating a Q-test statistic ($\chi^2$ – distributed) and I². The impact of trial quality, the behavioural techniques, and hours of training were examined by meta-regression. The impact of time since stroke, the shaping, and the control group were analysed by subgroup analysis, yielding a similar Q-test statistic. The variation of effect sizes within a subgroup was treated as random effects, while differences between sub-groups were treated as fixed effects. The study-to-study variance ($\chi^2$) was estimated as common for all subgroups. Funnel plots were used to evaluate signs of publication bias. Analysis of sensitivity included: (i) one-study removal analyses to detect potential outliers heavily influencing the results; (ii) inspection of forest plots; (iii) analysing single elements and different cut-offs of the adherence enhancing techniques; (iv) sub-group analysis and meta-regression of moderators in the sub-acute and chronic group to evaluate the effect of removing the heterogenic early intervention group.

RESULTS

Of 3,842 records initially screened 23 trials were included. The results from the literature research and selection of trials are shown in Fig. 1. Twelve papers reported supplemental analysis of another trial (16, 23–33). Three papers had insufficient data (34–36). Summary characteristics of the included trials are described in Table II. The mean age of the patients in the different trials ranged from 48.7 to 71.7 years. The proportion of women varied between 15% and 60%. Four trials were categorized in the early intervention group. The mean time post-stroke in this group ranged from 6 to 41 days. The mean time since stroke in 19 trials in the sub-acute and chronic group ranged from 75 days to 4.4 years. Twenty-one trials were classified as a RCT. Two trials were classified as a CCT, the first because of a non-randomized method of allocation (13) and the second because 11 patients who should have received the CIMT treatment were allocated to the control group and 10 controls received the CIMT treatment (37). For the latter study (37) the results from the per-protocol analysis were used. The PEDro score of the included trials ranged from 4 to 8 points. Regarding adherence enhancing techniques the home diary was most often reported ($n$= 8), followed by home practice ($n$= 18), and the use of MAL ($n$= 4), and the use of MAL ($n$= 4). A calibrated intensive group was used as control group in most studies ($n$= 18), followed by usual care ($n$= 3) and other control treatment ($n$= 2). Seven
### Table II. Characteristics of the included trials and classification of the moderators

<table>
<thead>
<tr>
<th>Study name</th>
<th>n/ Females (%)</th>
<th>Age, years, mean (SD)</th>
<th>Trial quality</th>
<th>Time since stroke</th>
<th>CIMT group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boake et al., 2007 (41) T: 9 C: 7</td>
<td>♀: 35%</td>
<td>T: 63.1 (14.3) C: 58.9 (14)</td>
<td>RCT PEDro: 6</td>
<td>Days 1–14</td>
<td>3 h/day (Shaping)</td>
<td>Intensive traditional therapy</td>
</tr>
<tr>
<td>Dahl et al., 2008 (39) T: 12 C: 12</td>
<td>♀: 42%</td>
<td>T: 62 (8) C: 60 (12)</td>
<td>RCT PEDro: 8</td>
<td>Mean days 700</td>
<td>Constraint 90%</td>
<td>3 h/day, 12 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Taub et al., 2012 (46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smania et al., 2008 (45) T: 9 C: 9</td>
<td>♀: 44%</td>
<td>T: 61.5 (13.7) C: 71.4 (5.3)</td>
<td>RCT PEDro: 5</td>
<td>Days 1–14</td>
<td>Constraint 90%</td>
<td>10 days (Usual care)</td>
</tr>
<tr>
<td>Lin et al., 2009 (50) T: 17 C: 17</td>
<td>♀: 60%</td>
<td>T: 64.5 (15.5) C: 64.7 (14.6)</td>
<td>RCT PEDro: 7</td>
<td>Mean days 10</td>
<td>Constraint 90%</td>
<td>(Calibrated intensive)</td>
</tr>
<tr>
<td>Khan et al., 2011 (40) T: 14 C: 14</td>
<td>♀: 37%</td>
<td>T: 60.4 (14.8) C: 60.4 (14.1)</td>
<td>RCT PEDro: 8</td>
<td>Mean days 75</td>
<td>Constraint 90%</td>
<td>3 h/day, 10 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Lin et al., 2007 (50) T: 17 C: 15</td>
<td>♀: 34%</td>
<td>T: 57.1 (18.3) C: 58.8 (15.2)</td>
<td>RCT PEDro: 6</td>
<td>Months 1–6</td>
<td>Constraint 6 h</td>
<td>10 days (Usual care)</td>
</tr>
<tr>
<td>Lin et al., 2009 (43) T: 16 C: 16</td>
<td>♀: 51%</td>
<td>T: 54.1 (11) C: 57.4 (12.8)</td>
<td>RCT PEDro: 7</td>
<td>Mean days 496</td>
<td>Constraint 90%</td>
<td>10 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Lin et al., 2010 (44) T: 5 C: 8</td>
<td>♀: 15%</td>
<td>T: 46.4 (26.0) C: 51.6 (12.4)</td>
<td>RCT PEDro: 4</td>
<td>Months 3–24</td>
<td>Constraint 90%</td>
<td>5 h constraint, 15 days</td>
</tr>
<tr>
<td>Myint et al., 2008 (45) T: 23 C: 20</td>
<td>♀: 58%</td>
<td>T: 63.4 (13.6) C: 63.9 (12.2)</td>
<td>RCT PEDro: 7</td>
<td>Mean days 41</td>
<td>Constraint 6 h</td>
<td>10 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Smania et al., 2012 (46) T: 30 C: 29</td>
<td>♀: 17%</td>
<td>T: 63.9 (9.56) C: 68.3 (12.68)</td>
<td>RCT PEDro: 8</td>
<td>Mean days 312</td>
<td>Constraint 90%</td>
<td>10 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Taub et al., 2006 (13) T: 21 C: 20</td>
<td>♀: 34%</td>
<td>T: 54.6 (12.1) C: 50.7 (19.2)</td>
<td>CCT PEDro: 4</td>
<td>Months &gt; 12</td>
<td>Constraint 90%</td>
<td>10 days (Usual care)</td>
</tr>
<tr>
<td>van der Lee et al., 1999 (37) T: 31 C: 31</td>
<td>♀: 44%</td>
<td>T: 59 (IQR 52–64) C: 62 (IQR 51–67)</td>
<td>CCT PEDro: 7</td>
<td>Months &gt; 12</td>
<td>Constraint 90%</td>
<td>10 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Wittenberg et al., 2003 (51) T: 9 C: 7</td>
<td>♀: 19%</td>
<td>T: 65 (RNG 41–81) C: 63 (RNG 50–75)</td>
<td>RCT PEDro: 6</td>
<td>Monthly 1,004</td>
<td>Constraint 90%</td>
<td>10 days (Usual care)</td>
</tr>
<tr>
<td>Wolf et al., 2006 (4) T: 98 C: 104</td>
<td>♀: 36%</td>
<td>T: 61.0 (13.5) C: 63.3 (12.6)</td>
<td>RCT PEDro: 6</td>
<td>Months 3–9</td>
<td>Constraint 90%</td>
<td>10 days (Usual care)</td>
</tr>
</tbody>
</table>
scales were classified as arm motor function measures, 9 as arm motor activity measures, 5 as ADL and 2 as measures of participation (Table I). The arm motor function data from one study (16) and the arm motor activity data from 2 studies (16, 38) were based on post-treatment data only. Ten studies reported follow-up data. We found 6-, 4-, and 3-month follow-up data in 2 studies (39, 40), one study (4) and 6 studies, respectively. Meanwhile, 1 study (37) reported data from 1.5 and 12 months of follow-up. A composite of these data were used to substitute for the 3–6 month effects in this study.

**Effect on arm motor function**

Sixteen trials (n = 714) used measures classified as arm motor function measures in the body function domain (4, 16, 19–22, 37, 38, 41–48). Fig. 2 shows a small effect size in favour of the CIMT groups (SMD 0.28, 95% CI 0.11–0.44). The heterogeneity statistic was not significant (I² = 11%, p = 0.331), and hence, none of the moderators interfered significantly with the results.

### Table II. Contd.

<table>
<thead>
<tr>
<th>Study name</th>
<th>n/ Females (%)</th>
<th>Age, years, mean (SD)</th>
<th>Trial quality</th>
<th>Time since stroke</th>
<th>CIMT group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu et al., 2007</td>
<td>T: 15 C: 15</td>
<td>T: 54.7 (8.6) C: 53.3 (6.3)</td>
<td>RCT PEDro: 7</td>
<td>Months 12–36</td>
<td>2 h/day (TP) Traditional rehabilitation</td>
<td>6 h constraint 2 h/day, 15 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Wu et al., 2007</td>
<td>T: 24 C: 23</td>
<td>T: 53.9 (11.2) C: 56.8 (12.9)</td>
<td>RCT PEDro: 6</td>
<td>3 weeks–36 months 6 h constraint</td>
<td>2 h/day (Shaping) Neurodevelopmental therapy</td>
<td>Mean days 373 2 h/day, 15 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Wu et al., 2007</td>
<td>T: 13 C: 13</td>
<td>T: 71.4 (6.4) C: 71.9 (6.8)</td>
<td>RCT PEDro: 6</td>
<td>Months 0.5–31</td>
<td>2 h/day (Shaping/TP) Traditional rehabilitation</td>
<td>Mean days 228 6 h constraint 2 h/day, 15 days (Calibrated intensive)</td>
</tr>
<tr>
<td>Wu et al., 2011</td>
<td>T1: 22 T2: 19 C: 18</td>
<td>T1: 51.9 (11.9) T2: 56.3 (12.2) C: 58.6 (11.6)</td>
<td>RCT PEDro: 6</td>
<td>Months &gt;6</td>
<td>2 h/day (TP) C1: Bimanual training</td>
<td>Mean days 477 6 h constraint C2: Control treatment</td>
</tr>
<tr>
<td>Wu et al., 2012</td>
<td>T1: 20 T2: 19 C: 18</td>
<td>T1: 54.0 (9.7) T2: 56.3 (12.2) C: 58.6 (11.6)</td>
<td>RCT PEDro: 6</td>
<td>Months 6–55</td>
<td>2 h/day (T1/T2: Shaping/TP, HD; T1: Trunk restraint) (Calibrated intensive)</td>
<td></td>
</tr>
<tr>
<td>Wu et al., 2012</td>
<td>T1: 15 T2: 15 C: 15</td>
<td>T1: 52.3 (11.3) T2: 54.9 (10.2) C: 54.3 (13.0)</td>
<td>RCT PEDro: 6</td>
<td>Months 6–59</td>
<td>2 h/day (Shaping) Neurodevelopmental therapy</td>
<td>Mean days 473 6 h constraint 2 h/day, 15 days</td>
</tr>
</tbody>
</table>

T: treatment group; C: control group; RCT: randomized controlled trial; CCT: controlled clinical trial; TP: task practice; BC: behavioural contract; HD: home diary; HSA: home assignments; MAL: motor activity log; ADL: activities of daily living.

**Fig. 2.** Forest plot of the random effects meta-analysis of constraint-induced movement therapy (CIMT) on arm motor function immediately after treatment. 95% CI: 95% confidence interval; ARAT: Action Research Arm Test; FMA: Fugl-Meyer Motor Assessment; Grippit: Grippit strength ratio; MESUPES: Motor Evaluation Scale for Arm in Stroke; WMFT: Wolf Motor Function Test; S&C: Subacute and chronic.
group was analysed alone the effect size decreased (SMD 0.22, 95% CI 0.06–0.40) and the statistical heterogeneity was zero. In terms of 3–6 months follow-up, 7 studies (n = 421) reported motor function outcomes (4, 16, 19, 37, 41, 45, 46). The effect size was small and not significant (SMD 0.12, 95% CI –0.36–0.60). The symmetrical funnel plot indicated no risk of publication bias.

**Effect on arm motor activity**

Twenty-two of the trials (n = 906) included measures of arm motor activities (4, 13, 16, 19–22, 37, 39–52). A moderate effect size was found for the effect of CIMT on arm motor activity (SMD 0.51, 95% CI 0.30–0.73). Fig. 3 shows the forest plot from the analysis. The heterogeneity was significant (I² = 66%, p < 0.001). The 2 trials classified as CCT (13, 37) produced high effect sizes, but the overall result was still significant when these trials were removed from the analysis (SMD 0.47, 95% CI 0.12–0.80). A moderator analysis showed that the effect size among the early intervention trials was negative and significantly smaller (SMD –0.15, 95% CI –0.77 to 0.47) than the other trials (SMD 0.61, 95% CI 0.38–0.83, Q = 5.2, p = 0.025). The heterogeneity in the effect was entirely explainable by the early intervention trials. When the sub-acute and chronic groups were examined alone, the I² statistic was 0% (p = 0.676). There was a slightly higher effect size in the 9 interventions that did not include shaping (SMD 0.64, 95% CI 0.30–0.99) although it was not significantly higher than in the other interventions (SMD 0.42, 95% CI 0.12–0.71, Q = 1.0, p = 0.323). There was no significant differences between the 18 trials that used a calibrated intensive control group and those 5 that controlled for usual care or other control treatment (Q = 0.03, p = 0.872). Meta-regression did not show any significant effect of the PEDro score, the number of adherence techniques, or the duration of treatment. No signs of publication bias were found in the funnel plots.

Nine trials reported follow-up data within the 3–6 month range for 478 patients (4, 16, 19, 37–41, 45, 46). A moderate effect size was reported on arm motor activity (SMD 0.41, 95% CI 0.11–0.70). Considerable heterogeneity was observed between trials (I² = 68%, p = 0.001). Following removal of the trial not using proper randomization (37), the effect size was no longer significant (SMD 0.33, 95% CI –0.02 to 0.67). A trend toward lower effect sizes was also evident in the early intervention trials (SMD 0.13, 95% CI –0.50–0.79), but the trend was not significant (Q = 0.88, p = 0.348). At this point, significant heterogeneity was seen among both the early (I² = 75%, p = 0.017) and the sub-acute and chronic groups (I² = 68%, p = 0.007).

**Effects on activities of daily living and participation**

Ten trials (19, 21, 37, 39, 42, 43, 45, 48, 50, 52) reported the effects for 334 participants in the other activity measurements. The results from analysis of these trials are shown in Fig. 4. The effect size was smaller and not significant (SMD 0.019, 95% CI –0.02–0.39). No statistical heterogeneity was observed (I² = 0%, p = 0.840) and, hence, none of the moderators influenced the effect size significantly. Only 2 studies (37, 39) reported follow-up data on the ADL (n = 90). The combined SMD was 0.06 (95% CI –0.33–0.45). Two studies (21, 50) reported post-treatment data for the participation category. A moderate, but non-significant, effect size was found (SMD...
839

Systematic review and meta-analysis of CIMT

0.23, 95% CI –0.24–0.71). Only one of these studies reported follow-up data.

DISCUSSION

This meta-analysis found a small, but significant, effect of CIMT on arm motor function immediately after treatment in patients with stroke. A moderate effect on arm motor activity was also found immediately after treatment and at 3–6 months follow-up, although this effect was dependent on the inclusion of one non-randomized trial. A large heterogeneity was observed among the studies in the arm motor activity domain, indicating that one or more factors could affect the treatment.

The heterogeneity was predominately present in the rather few studies of early interventions. None of the other moderators (i.e. the quality of the trial, adherence-enhancing activities, shaping, the control group, or the duration of training) interfered significantly with the effect sizes or were able to explain the heterogeneity in our material. Despite a low number of follow-up studies, our results do not anticipate lasting effects on arm motor function. The number of studies was too low to draw conclusions about the effect on ADL and participation.

Our systematic meta-analysis included 23 studies and reported the synthesized effect sizes on up to 22 trials examining up to 906 stroke patients. The largest former meta-analysis is a Cochrane review with a subsequent update (7, 10). Compared with our study they included less intensive CI interventions (less than 2 h of training per day), as well fewer trials (n = 14) and fewer patients (n = 477). Of the 17 studies in our arm motor activity analysis, only 3 were included in the former meta-analysis (7). Of the 23 studies in our arm motor activity analysis, only 6 were included in former analyses of arm motor function (10). Shi et al. (9) analysed CIMT lasting 0.5–3 h and constraint use of <6 h, where the effect was compared with traditional rehabilitation.

This provides a slightly different analysis from the current one, which included interventions lasting from 2 to 7 h. In addition, most of the trials included in our review compared CIMT with an equally intensive treatment that would not automatically fit the description of traditional rehabilitation used by Shi.

Our effect size on arm motor function is lower than that reported by Sirtori et al. (7) (SMD 0.65), which might be explained by greater number of trials and the different measurements included. In the arm motor activity domain our effect sizes are higher than reported by Corbetta et al. (10) on arm motor function (SMD 0.44) and lower than reported on perceived arm use by Sirtori et al. (7) (SMD 1.16) and Paura et al. (8) (SMD 0.85). Our arm motor activity domain included both these categories and this is probably the reason for the intermediate result. Our effect sizes on ADL immediately after treatment are similar to those calculated by Corbetta et al. (10) for the effect on disability (SMD 0.21, 95% CI –0.06 to 0.65). Meanwhile, their follow-up results were similar to ours. For the Functional Independence Measure, Peurala et al. (8) found a non-significant difference of 0.08 in 9 trials that included 267 patients. This result is similar to our results for ADL.

One reason for the lower effect sizes in our analysis compared with that of other analyses could be that we standardized the mean difference by the post-treatment standard deviation (SD). This decision was made a priori because the post-treatment SD was thought to best reflect the variance at the time when the effect was evaluated. Some of the other studies have been standardized by the change-score SDs (7, 9), while other studies have mixed standardization by post-test SDs and change-score SDs (8). The pre-treatment variation was often smaller and could therefore yield larger effect sizes, which might be the reason for the lower SMDs in our study. Another reason for lower effect sizes might be our pre-specified inclusion criteria that omitted a common CIMT modification that use 0.5 h of therapy per day, and number of treatment days per week was chosen to lower the clinical heterogeneity between studies, but including these studies might have increased the effect sizes and the dose–response relationship between hours of treatment and treatment results.

A higher effect size was expected in trials with extensive use of behavioural techniques; however, this hypothesis was not supported by the current meta-analysis. A small experimental study has shown that the use of behavioural techniques may have
a more long-lasting effect than treatments without them (53). Due to the low power of the follow-up analyses, we might have missed such an effect in our material, and this will be a concern for future research. In addition, some challenges were encountered in the sampling of the moderator variables. Comparing our definitions of shaping and adherence-enhancing activities with the descriptions in the reports was difficult; hence, the chance that behavioural techniques were misclassified is possible. However, research milieus known to emphasize these techniques describe them in detail (4, 13); thus, despite the lack of information sometimes observed, we believe that the behavioural techniques were probably correctly classified in most cases.

The present meta-analysis raised concerns regarding the large variation in treatment effects among the early intervention trials. One of the early interventions trials (19) can be considered an outlier, which also contributed significantly to the heterogeneity between trials. The study had a 3-armed design with 2 treatment groups, which included patients very early after stroke. In terms of arm motor function, the study found that the Action Research Arm Test scores in the 2-h CIMT group were positively affected by the treatment, whereas those in the 3-h CIMT group were substantially negatively affected compared with controls (19). This study used a different measurement (Stroke Impairment Scale – hand) than the other studies, and the pre-treatment scores for this measurement were not reported. As such, this may be one of the reasons for the large negative effect on arm motor activity shown in this study only. The variation in effect sizes between the early intervention trials could indicate that the results of CIMT treatment at this stage may be more sensitive to intensity, other treatment components, or individual factors (54). Future studies on the efficacy of CIMT should therefore specifically compare early with later interventions in order to settle this question.

The new insight from our analyses is that there is a consistent effect of CIMT on arm motor function and arm motor activities in sub-acute and chronic patients. These studies represent a large variation in treatment dose, age groups, gender distribution, countries, and control groups. Nonetheless, the effect in the sub-acute and chronic group was still uniform. Even though the variation of effect sizes between the trials might be clinically significant, this variation could be expected based on the sizes of the trials. Because of our definition of the sub-acute and chronic group (mean time from stroke >45 days) this group will also include some patients included between 2 weeks and 3 months post-stroke, not usually classified as sub-acute or chronic. Even though this interferes with the external validity the stable effect in this group does not suggest large implications from this classification. The post-treatment effect on arm motor activity was larger than the effect on arm motor function. CIMT have been developed to overcome learned non-use and increase the use of the more affected arm (2, 3), which probably explains why we observed a larger effect on arm motor activity measures compared with arm motor function measures. Another explanation may be that the measurement most often included in our arm motor activity category (i.e. MAL) was specifically designed to measure the effect of CIMT, and may therefore be more sensitive to the treatment effect (55). Our analysis also showed a moderate effect on arm motor activity 3–6 months post-stroke; a result that has not been shown by former meta-analyses. Future studies should aim at explaining the heterogeneity in the follow-up results.

**Conclusion**

By combining measurements in the ICF domains, we were able to show significant post-treatment effects of CIMT on arm motor function and arm motor activity and a follow-up effect on arm motor activity. In particular, stable effects immediately after treatment were detected in the sub-acute and chronic group. Future research should focus on the clinical and societal importance of these effects. More research is also needed to determine the effect of moderators in the treatment. We are concerned about the limited data and unpredictable effect from early intervention trials. We recommend that experiments with CIMT in the early stage after stroke be conducted with care. Significant emphasis should be placed on explaining the components that contribute to favourable or non-favourable effects from the treatment at this stage.

**ACKNOWLEDGEMENTS**

The authors would like to thank Bjørn Eivind Kirsebom and Martin Vatshaug for participating in the literature review and data extraction for this study. The Northern Norway Regional Health Authority, Helse Nord RHF, funded parts of the study (No 3228).

**REFERENCES**

9. Shi YX, Tian JH, Yang KH, Zhao Y. Modified constraint-induced movement therapy versus traditional rehabilitation in patients with...