Effects of Robot-Assisted Therapy on Upper Limb Recovery After Stroke: A Systematic Review

Gert Kwakkel, PhD, Boudewijn J. Kollen, PhD, and Hermano I. Krebs, PhD

Objective. The aim of the study was to present a systematic review of studies that investigate the effects of robot-assisted therapy on motor and functional recovery in patients with stroke. Methods. A database of articles published up to October 2006 was compiled using the following Medline key words: cerebrovascular accident, cerebral vascular disorders, stroke, paresis, hemiplegia, upper extremity, arm, and robot. References listed in relevant publications were also screened. Studies that satisfied the following selection criteria were included: (1) patients were diagnosed with cerebral vascular accident; (2) effects of robot-assisted therapy for the upper limb were investigated; (3) the outcome was measured in terms of motor and/or functional recovery of the upper paretic limb; and (4) the study was a randomized clinical trial (RCT). For each outcome measure, the estimated effect size (ES) and the summary effect size (SES) expressed in standard deviation units (SDU) were calculated for motor recovery and functional ability (activities of daily living [ADLs]) using fixed and random effect models. Ten studies, involving 218 patients, were included in the synthesis. Their methodological quality ranged from 4 to 8 on a (maximum) 10-point scale. Results. Meta-analysis showed a nonsignificant heterogeneous SES in terms of upper limb motor recovery. Sensitivity analysis of studies involving only shoulder–elbow robotics subsequently demonstrated a significant homogeneous SES for motor recovery of the upper paretic limb. No significant SES was observed for functional ability (ADL). Conclusion. As a result of marked heterogeneity in studies between distal and proximal arm robotics, no overall significant effect in favor of robot-assisted therapy was found in the present meta-analysis. However, subsequent sensitivity analysis showed a significant improvement in upper limb motor function after stroke for upper arm robotics. No significant improvement was found in ADL function. However, the administered ADL scales in the reviewed studies fail to adequately reflect recovery of the paretic upper limb, whereas valid instruments that measure outcome of dexterity of the paretic arm and hand are mostly absent in selected studies. Future research into the effects of robot-assisted therapy should therefore distinguish between upper and lower robotics arm training and concentrate on kinematic analysis to differentiate between genuine upper limb motor recovery and functional recovery due to compensation strategies by proximal control of the trunk and upper limb.

Key Words: Robotics—Cerebrovascular accident—Activities of daily living—Upper limb—Review—Systematic.

Stroke is the leading cause of disability in the United States. A total of 750000 individuals are affected each year and the prevalence rate is about 200 to 300 patients per 100 000 inhabitants.1 Although prospective epidemiological studies are lacking, findings of several longitudinal studies indicate that in 30% to 66% of hemiplegic stroke patients, the paretic arm remains without function when measured 6 months after stroke, whereas only 5% to 20% demonstrate complete functional recovery.

The results of a systematic review involving 123 randomized clinical trials (RCTs) by van Peppen and colleagues demonstrated that there is strong evidence that intensity as well as task specificity are the main drivers in an effective treatment program after stroke. In addition, this training should be repetitive, functional, meaningful, and challenging for a patient.2,4

However, the question as to how the effects of exercise therapy can be further enhanced in a clinical environment presents a challenge to answer. Therefore, there is a need to develop better ways to augment exercise training in a functional way. Using therapeutic adjuncts to facilitate clinical practice, such as robotics,5,11 is a new promising development. Robotics allows patients to train independently of a therapist and to improve on their own functional level (ie, robot-assisted therapy). In particular, there is strong evidence for robot-assisted therapy to increase treatment compliance by way of introducing incentives to the patient, such as games. In addition, by using computer-assisted devices for

DOI: 10.1177/15459683070305457

Copyright © 2008 The American Society of Neurorehabilitation

111
regaining upper limb function, the robot can easily apply new constraints, to optimize the required movement pattern. Therefore, the complexity of a motor task to be learned can be controlled far more precisely with robotics than in conventional treatment approaches.

Although many devices have been designed to deliver arm therapy in individuals with stroke, 5 of these devices, the MIT-MANUS 6,7 (designed and built at the Massachusetts Institute of Technology; Interactive Motion Technologies, Inc, Cambridge, MA), the ARM Guide 8 (Assisted Rehabilitation and Measurement guide), the MIME 9 (Mirror-Image Motion Enabler), the InMotion 2 Shoulder-Elbow Robot, 10 and the Bi-Manu-Track 11 were tested in at least 1 RCT. The MIT-MANUS is a robot that allows subjects to execute reaching movements in the horizontal plane. This 2 degrees of freedom (DoF) robot enables unrestricted movements of the shoulder and elbow joints. 9 The ARM Guide is a trombone-like device and has 4 controlled DoF. A DC servo motor can assist in the movement of a subject's arm in the reaching direction along a linear track. Optical encoders record the position in the reach, elevation, and yaw axes. 7 The MIME robot consists of a 6-DoF robot arm. The robot enables the bilateral practice of a 3-DoF shoulder-elbow movement, whereby the nonparetic arm guides the paretic arm. 12 The InMotion 2 Shoulder-Elbow Robot, which is the commercial version of MIT-MANUS, has 2 DoF and provides shoulder-elbow training in the horizontal plane with a supported forearm. 10 The Bi-Manu-Track is designed to specifically train distal arm movements by practicing bilateral elbow pronation and supination as well as wrist flexion and extension in a mirror or parallel fashion. 11

In the past, several studies were unable to prove superiority of 1 type of conventional stroke regimen over another, 13,14 but there is strong evidence that highly repetitive movement training can result in improved recovery. 10,15 Applying robot-assisted therapy enables patients to practice intensively with their upper paretic limb. The objective of the present systematic review is to determine the additional effects of robot-assisted therapy on motor recovery and functional outcome in comparison with conventional treatment forms.

METHODS

Literature Search

A computerized literature search was conducted in Medline, CINAHL, EMBASE, Cochrane Controlled Trials Register, DARE, SciSearch, DocOnline, and PEDro. Studies were collected up to October 2006. MeSH keywords used were: cerebral vascular accident, cerebral vascular disorders, stroke, paresis, hemiplegia, upper extremity, arm, and robot (ie, all word combinations that start with the term "robot"). Literature lists of narrative reviews were also evaluated for relevant publications. Only articles written in the English, German, or Dutch language were included. Studies were included when: (1) patients were diagnosed with cerebral vascular accident; (2) effects of robot-assisted therapy for the upper limb were investigated; (3) the outcome was measured in terms of motor and/or functional recovery of the upper paretic limb; and (4) the study was a randomized clinical trial (RCT). Excluded were studies that compared the effects of 2 different types of robot-assisted therapy and studies of persons with chronic impairment due to stroke that compared discharge outcomes with preintervention stable scores.

Definitions

Cerebral vascular accident has been defined as "a sudden, non-convulsive loss of neurologic function due to an ischemic or hemorrhagic intracranial vascular event" (PubMed [Medline], MeSH database, 2005). Robotics has been defined as: "The application of electronic, computerized control systems to mechanical devices designed to perform human functions." Although formerly restricted to the industry, nowadays certain human functions can also be controlled by bionic (bioelectronic) devices, such as automatic insulin pumps and other prostheses (PubMed [Medline], MeSH database, 2005). One independent reviewer (JH) selected articles based on title and abstract.

Methodological Quality

The methodological quality of the studies was rated with the PEDro scale 17 and scored by 2 independent reviewers (JH and GK). Interrater reliability of individual items was tested with Cohen's kappa statistics. When no consensus between the 2 reviewers was reached, a third reviewer made the final decision. Reviewers were not blinded to author(s), institution(s), or journal. PEDro scores of 4 points or more were classified as "high quality," whereas studies with 3 points or fewer were classified as "low quality." 17

Quantitative Analysis

The abstracted data (mean age, numbers of patients in the experimental and control groups, mean difference-in-change scores on the measure scales of motor recovery and functional level, and their standard deviations in experimental and control groups at baseline) were
entered into Excel for Windows. The effect size \(g\) (Hedges’s \(g\)) for individual studies was established by calculating the difference between the means of the experimental and control groups divided by the average population standard deviation (SD). If necessary, means and SD, were requested from the respective authors. Alternatively, Hedges’s \(g\) estimates were obtained from \(t\) values. The impact of sample size was determined by assigning a weighting factor \((w_i)\) to each study, in such a way that larger effect weights were given to studies with larger samples. Finally, \(g^2\) values of individual studies were averaged, resulting in a weighted summary effect size (SES), and the individual weights were combined to estimate the variance of the SES. The effect size \(g^2\) for individual studies was computed for the degree of motor recovery and functional performance. If a significant between-study variation (Q-statistic) was found (representing statistical heterogeneity) a random effects model was applied. Based on the classification of Cohen, effect sizes \(<0.2\) were classified as small, between \(0.2\) and \(0.5\) as medium, and \(>0.5\) as large. For all outcome variables, the critical value for rejection of the \(H_0\) hypothesis was set at \(P < .05\) (2-tailed).

**RESULTS**

Appendix A shows the flow chart for the selection of studies. After searching the electronic databases 87 from 173 hits were considered to be relevant for further screening. After screening their abstracts, however, from these 87 publications, 44 relevant studies were selected. Ten of these articles were critical or narrative reviews, and 24 studies were not controlled trials. Following the exclusion of (1) preexperimental studies and (2) controlled studies that did not measure motor and/or functional recovery of the upper paretic limb, in total 22 studies were excluded. Finally, 10 studies were identified as being relevant. Two studies referred to the same patient sample. The study by Volpe et al. presented outcomes on motor recovery of the upper paretic limb (Fugl-Meyer arm motor score [FMA]), whereas the study of Fasoli et al. reported on outcome of functional ability (functional independence measure [FIM]) in the same stroke population.

Table 1 shows the main characteristics of the 10 eligible studies included in the present meta-analysis. The start of the therapy ranged from 1 week after stroke to more than 6 months after stroke.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Participants</th>
<th>Score</th>
<th>Effect Size (95% CI)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen</td>
<td>2007</td>
<td>N=31</td>
<td></td>
<td>0.54 (0.46 - 0.61)</td>
<td></td>
</tr>
<tr>
<td>Burge</td>
<td>2006</td>
<td>N=29</td>
<td></td>
<td>0.50 (0.43 - 0.57)</td>
<td></td>
</tr>
<tr>
<td>Kahn</td>
<td>2005</td>
<td>N=19</td>
<td></td>
<td>0.50 (0.36 - 0.65)</td>
<td></td>
</tr>
<tr>
<td>Volpe</td>
<td>2005</td>
<td>N=56</td>
<td></td>
<td>0.50 (0.36 - 0.65)</td>
<td></td>
</tr>
<tr>
<td>Lum</td>
<td>2002</td>
<td>N=27</td>
<td></td>
<td>0.50 (0.36 - 0.65)</td>
<td></td>
</tr>
<tr>
<td>Hesse</td>
<td>2002</td>
<td>N=35</td>
<td></td>
<td>0.50 (0.36 - 0.65)</td>
<td></td>
</tr>
<tr>
<td>Daily</td>
<td>2006</td>
<td>N=12</td>
<td></td>
<td>0.50 (0.36 - 0.65)</td>
<td></td>
</tr>
<tr>
<td>Kahn</td>
<td>2006</td>
<td>N=13</td>
<td></td>
<td>0.50 (0.36 - 0.65)</td>
<td></td>
</tr>
<tr>
<td>Lum</td>
<td>2006</td>
<td>N=15</td>
<td></td>
<td>0.50 (0.36 - 0.65)</td>
<td></td>
</tr>
</tbody>
</table>

**S.E.S.** (random effects model)

Favors Control 

Favors Treatment

Figure 1. Meta-analysis of robot-assisted therapy trials on motor recovery. SES = summary effect size.

agreement was 0.82. The methodological quality varied from 4 points to 8 points.

The treatment arms of 2 studies were not comparable at baseline. In the study by Hesse et al., the Barthel Index was statistically significantly higher in favor of the experimental group. The study by Kahn et al. showed a statistically significantly higher value for the Chedoke-McMaster test in favor of the control group at baseline.

**Quantitative Analysis**

**Motor recovery.** Seven studies used the F-M as outcome parameter, whereas the studies of Kahn et al. evaluated outcome with the Chedoke-McMaster. A total of 218 patients with stroke were involved. Five RCTs reported statistically significant effects for motor recovery in favor of the experimental group, whereas 4 RCTs did not find significant differences. On average, the experimental group received daily 48.3 minutes of robot-assisted therapy (RT) and the control group 29.0 minutes of control therapy (CT). An overall statistically nonsignificant (0.65, 95% confidence interval [CI] 0.02 to 1.33; \(Z = 1.90, P = .06\)) heterogeneous SES (\(\chi^2 = 40.82, P < .001\)) was found in favor of the robot-assisted therapy (Figure 1 and Table 2). As shown in Table 2 and Figure 1, observed heterogeneity between studies was mainly due to a study of Hesse and colleagues that reported a larger individual effect size for bilateral distal arm training when compared with the other studies. Subsequent sensitivity analysis without the study of Hesse et al. showed a homogeneous SES (\(\chi^2 = 4.35, P = .60\)) in favor of shoulder-
<table>
<thead>
<tr>
<th>Reference</th>
<th>Stroke Type</th>
<th>Severity (F-M [U-L] at Baseline)</th>
<th>Start of RT/CT (E/C)*</th>
<th>Type of Intervention (E/C)</th>
<th>Intervention Categories</th>
<th>Daily RT (E/C)b</th>
<th>Daily CT (E/C)b</th>
<th>Mean Age (years) (E/C)a</th>
<th>Outcome</th>
<th>Author Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aisen 1997</td>
<td>Hemorrhagic, ischemic</td>
<td>13.8/17.1</td>
<td>2.8/3.3 wk</td>
<td>RT vs. robot exposure (control)</td>
<td>MIT-MANUS</td>
<td>60/0</td>
<td>±0/10</td>
<td>58.5/63.3</td>
<td>F-M FIM</td>
<td>Significant difference in motor recovery (acute patients)</td>
</tr>
<tr>
<td>Burgar 2000</td>
<td>All types</td>
<td>24.8/21.8</td>
<td>26.5/26.4 mo</td>
<td>RT vs. neuro-developmental therapy</td>
<td>MIME</td>
<td>36/0</td>
<td>0/36</td>
<td>64.4/63.3</td>
<td>F-M FIM</td>
<td>Significant difference in motor recovery (chronic patients)</td>
</tr>
<tr>
<td>Kahn 2000</td>
<td>?</td>
<td>?</td>
<td>&gt;6 mo</td>
<td>RT vs. unassisted, unrestrained reaching exercises</td>
<td>ALM Guide</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>Ch McM</td>
<td>Repetitive movements seem to be the primary stimuli to recovery</td>
</tr>
<tr>
<td>Volpe 2000</td>
<td>Hemorrhagic, ischemic</td>
<td>8.6/10.5</td>
<td>22.5/26.0 days</td>
<td>RT vs. robot exposure (control)</td>
<td>MIT-MANUS</td>
<td>60/0</td>
<td>0/12</td>
<td>62/67</td>
<td>F-M</td>
<td>Improvement of the motor performance of the exercised shoulder and elbow</td>
</tr>
<tr>
<td>Fasoli 2004</td>
<td>Hemorrhagic, ischemic</td>
<td>8.6/10.5</td>
<td>9/10 days</td>
<td>RT vs. robot exposure (control)</td>
<td>MIT-MANUS</td>
<td>60/0</td>
<td>0/12</td>
<td>62/67</td>
<td>F-M</td>
<td>Intensive therapy leads to better recovery after stroke</td>
</tr>
<tr>
<td>Lum 2002</td>
<td>All types</td>
<td>24.8/26.6</td>
<td>30.2/28.8 mo</td>
<td>RT vs. neuro-developmental therapy</td>
<td>MIME</td>
<td>36/0</td>
<td>0/36</td>
<td>63.2/65.9</td>
<td>F-M FIM</td>
<td>Significant difference in motor recovery</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Reference</th>
<th>Stroke Type</th>
<th>Severity (F-M [U-L] at Baseline)</th>
<th>Start of RT/CT (E/C)</th>
<th>Type of Intervention (E/C)</th>
<th>Intervention Categories</th>
<th>Daily (min) RT (E/C)</th>
<th>Daily (min) CT (E/C)</th>
<th>Mean Age (years) (E/C)</th>
<th>Outcome</th>
<th>Author Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hesse 2005</td>
<td>Ischemic (first stroke)</td>
<td>7.9/7.3</td>
<td>4-8 wk&lt;sup&gt;c&lt;/sup&gt;</td>
<td>RT vs. electrical stimulation</td>
<td>Bi-Manu-Track</td>
<td>20/0</td>
<td>0/20</td>
<td>65.4/64.0</td>
<td>F-M</td>
<td>Superior improvement in upper limb control and power</td>
</tr>
<tr>
<td>Daly 2005</td>
<td>Hemorrhagic, ischemic</td>
<td>21/23</td>
<td>&gt;12 mo</td>
<td>RT vs. functional neuromuscular stimulation and motor earning</td>
<td>InMotion&lt;sup&gt;2&lt;/sup&gt; Shoulder-Elbow Robot</td>
<td>90/0</td>
<td>0/90</td>
<td>21-62</td>
<td>F-M</td>
<td>Significant gains in F-M upper-limb coordination</td>
</tr>
<tr>
<td>Kahn 2006</td>
<td>?</td>
<td>3.5/3.2</td>
<td>75.8/103.1 mo</td>
<td>RT vs. task-matched amount of unassisted reaching</td>
<td>ARM Guide</td>
<td>61/0</td>
<td>0/61</td>
<td>55.6/55.9</td>
<td>Ch McM</td>
<td>Robotically assisting in reaching successfully improved arm movement ability</td>
</tr>
<tr>
<td>Lum 2006</td>
<td>?</td>
<td>8.4/5.0</td>
<td>10.0/10.6 wk</td>
<td>RT vs. conventional therapy</td>
<td>MIME</td>
<td>45/0</td>
<td>0/45</td>
<td>69.8/59.9</td>
<td>F-M, FIM</td>
<td>Robot-assisted treatment gains exceeded those expected from spontaneous recovery</td>
</tr>
</tbody>
</table>

Abbreviations: F-M [U-L] = Fugl-Meyer (upper-limb); RT = robot-assisted therapy; CT = control therapy; min = minutes; E/C = experimental group vs. control group; MIME = Mirror-Image Motion Enabler; ARM Guide = Assisted Rehabilitation and Measurement guide; wk = week(s); mo = month(s); FIM = Functional Independence Measure; Ch McM = Chedoke-McMaster; ? = not specified or unknown.

<sup>a</sup> Only median figures given.

<sup>b</sup> Average of calculated minutes for every working day during intervention.

<sup>c</sup> Average age of experimental and control group together.
elbow arm robotics (0.36 SDU [fixed]; 95% CI 0.05 to 0.65; Z = 2.32, P < .026).

Activities of Daily Living (ADLs)

Five studies\textsuperscript{a,b,d4,f5} evaluated outcome with the FIM. A total of 139 patients with stroke were involved. None of the studies reported significant effects for ADL in favor of the experimental group. The study by Burgar et al (2000) is not shown due to the inability to calculate an individual effect size. On average, the experimental group received daily 50.3 minutes of RT and the control group 23.4 minutes of CT. A homogeneous nonsignificant SES was found ($\chi^2 = 0.50$, $P > .05$) for robot-assisted therapy (SES [fixed] 0.13 SDU; 95% CI: $-0.23$ to 0.50; $Z = 0.86$, $P > .05$) (Figure 2 and Table 3).

DISCUSSION

The present review involving 218 patients shows a positive trend toward robot-assisted therapy for the proximal upper limb when compared with conventional treatment modalities with regard to motor recovery when measured with the FMA.
Effects of Robot-Assisted Therapy on Upper Limb Recovery After Stroke

No significant improvements were observed on outcomes of ADLs measured by the FIM scale. It is common knowledge that the FIM (like the Barthel Index) does not measure dexterity of the upper paretic limb properly. Therefore, scales that measure dexterity specifically, such as the Action Research Arm test, the Wolf Motor Function Test, and the Jebsen Test or Nine-Hole Peg Test, are preferred. In addition, as was shown recently by Rohrer et al., and Kahn and colleagues, future studies should also investigate kinematic changes to better understand the observed improvements in motor coordination. For this later purpose, longitudinally conducted studies with repeated measurements are needed to understand how synergistic-dependent movement patterns improve in patients during robotic assisted therapies. In particular, changes that occur in adaptive compensatory movements of the trunk should be investigated. Moreover, it should be noted that measurements at a functional level cannot differentiate between improvements at the motor recovery level and improvements due to the use of alternative compensating strategies. Therefore, a better understanding of the relationship between the mechanisms of cortico-motor organization and motor recovery is needed. In particular, understanding the bihemispheric plasticity after stroke will allow the development of new (robot-assisted) therapies to enhance learning-dependent neuroplasticity. Although an initial small study by Lunn and colleagues suggested that unilateral training is superior to bilateral training, larger studies investigating whether bilateral robotic assisted training, with or without auditory or visual cuing and feedback that promotes inter-hemispheric activation of the limbs, should be subject to investigation. It may have an added value to only unilateral training, particularly to the more impaired patients. New developments in robotic innovations and capabilities are still to be expected and most likely will expand its applicability to other or more specific motor functions.

Another important issue to elucidate is the impact of high intensity robot-assisted therapy on motor recovery after stroke. In studies with the MIT-MANUS robot, in addition to conventional therapy, patients received 5 hours per week of robot-assisted therapy, whereas the control group received only 1 hour of robot exposure. The number of movements generated in robot-assisted therapy is far higher than in other forms of therapy, such as electric stimulation (ES), free reaching, and neurodevelopmental therapy (NDT). It can be concluded that high intensity repetitive movements constitute an important contributor to the effectiveness of robot-assisted therapy. In fact, studies that tried to match the intensity of robotic therapy to the number of movements generated by other forms of therapy failed to show a differential effect. In other words, robotic therapy had no particular advantage at low utilization, but it also did not hinder or halt recovery.

the arm and hand impairment part of the Chedoke-McMaster Stroke Assessment Scale (CMSA). The lack of overall significance for all included studies on motor recovery was the result of the application of a random effects model (REM). This REM calculates a summary effect size (SES) that is based on the within- and between-study variance. As a consequence, larger 95% confidence intervals of the SES are obtained than a fixed effects model would generate. However, a sensitivity analysis showed that observed heterogeneity was mainly caused by the study of Hesse and colleagues, who investigated the effects of the Bi-Manu-Track robotic on distal elbow pro-supination and wrist flexion and extension activities in patients with stroke. By contrast, the other robots, such as the MIT-MANUS and MIME robots, are designed to train proximal shoulder and elbow movements. This finding suggests that the observed effect size is dependent on proximal or distal arm robot training. Unfortunately, only 3 of the 7 studies that measured the FMA or CMSA score provided additional subscores for the proximal and distal arm components at baseline and at the end of the therapy. Two studies reported only the subscores of both components at the end of the therapy, whereas the other studies did not differentiate between these scores at all. Therefore, it remains uncertain whether the observed improved motor score is due to an improvement at the proximal shoulder-elbow level or at the distal hand-wrist level. To better understand what exactly patients learn by robotics when they improve in upper limb motor function, future studies should address this issue. The significant, moderate SES of upper arm robotics on motor control based on FMA and CMSA scales denotes a mean overall change of 7% to 8% in motor control of the upper limb in favor of the robot-assisted therapy. These effects were based on studies of high methodological quality (PEDro score = 6 ± 1.25 [mean ± SD]).

Neurorehabilitation and Neural Repair 22(2); 2008
To examine whether the effectiveness of robot-assisted therapy is due to the treatment modality or to the high intensity of training, dose-response trials are required. It is important to understand that robotic therapy simply uses robots as vehicles to deliver highly repetitive therapy. There is no reason to assume that robots will lead to better results than human-delivered movement therapy if there is a match of all variables. In fact, 1 of us (HIK) has just shown that in a small study (Volpe BT, Lynch D, Ferraro M, Galgano M, Hogan N, Krebs HI, unpublished data). In practice the high intensity of training in patients receiving robotics in the clinic in a classroom format or also in their own home environment should be subject of cost-effectiveness studies. These studies should incorporate the purchase of a robotic device and support with a single therapist supervising multiple patients, as demonstrated by Daly and colleagues, and the costs of intermittent, patient-tailored training by a therapist in a clinical setting. For the home environment, patients’ performances can be monitored and data collected remotely using robotic therapy in conjunction with broadband telematics. This development has—among other benefits—the potential of saving traveling expenses to and from a health care facility and increasing of independent training time. To date, high quality studies investigating the efficiency of robotics in relation to usual care are lacking in the literature. Such trials should incorporate a critical and comprehensive economic and effectiveness evaluation of hospital and community care, patient and family resources, and other resources, for example, home help visits, and involve a cost-effectiveness analysis rather than a cost-minimization, cost-benefit, or cost-utility analysis. To our knowledge, the only such a study presently in progress is being supported and run by the Veterans Affairs in the United States, involving 4 of its rehabilitation hospitals (PI: Albert Lo, Yale University and VA West Haven; CSP 558; West Haven, CT; Baltimore, MD; Gainesville, FL; Seattle, WA).

Finally, most studies were based on a small sample size and were heavily underpowered to reject the H0 hypothesis. In particular, because the stroke population is heterogeneous and treatment effects are relatively small when compared with observed differences in patterns of recovery due to spontaneous recovery, stratification on the basis of prognostic baseline characteristics becomes necessary. For example, assuming that a 6-point difference constitutes the minimal clinically important difference (MCID) on outcome of ARAT, at least 447 patients per arm are needed in a 2-arm RCT to maintain 80% statistical power to reject the H0 hypothesis. The large standard deviation of ARAT measurements, that is, about 32 points at 26 weeks after stroke, is mainly responsible for this large sample size. In contrast, when selecting only stroke subjects displaying some dexterity on the ARAT (ie, ARAT >9 points), a sample standard deviation of 12 points is observed and only 63 patients in each arm will be needed to show an MCID of 6 points. As is the case in constraint-induced movement therapy (CIMT), demonstrating a therapeutic effect is a matter of selection of appropriate patients.

The present review has a number of limitations. First, in this review we assumed that all studies used different patients, but because some studies were conducted at the same time and at the same place, we cannot be certain that only unrelated study populations were used. Second, in the present study we pooled Fugl-Meyer arm motor scores with the scores from the physical impairment inventory of the arm of the Chedoke-McMaster Stroke Assessment Scale to calculate 1 overall summary effect size for motor outcome of the upper paretic limb. The FMA includes 42 items to measure shoulder-elbow impairment and coordination and 24 items to measure the hand-wrist impairment. On the other hand, the CSMA contains 2 dimensions related with motor control of arm and hand each measured on a 7-point scale. In our opinion, pooling outcomes of both scales to generate 1 dimensionless SES is justified because both assessments are based on the 6 recovery stages of Brunnstrom. This suggests that both scales measure the same underlying construct (ie, synergy dependency of motor control). Third, we also pooled the robot-assisted interventions to obtain a overall effect size, although different robots focused on different parts of the upper limb function, different robot control strategies were used (robots can be programmed to deliver different behaviors and, as stated earlier, we excluded comparison of different types of robot-assisted therapy), and some protocols focused on the subacute phase of stroke recovery whereas other studies focused on the chronic phase. For the same reason we were unable to differentiate between the control interventions. Different types of control interventions, such as NDT and ES, which have proven their effectiveness only at the impairment level, made it difficult to interpret the effectiveness of robot-assisted therapy. These shortcomings emphasize the need for more high quality RCTs. In the present review we did pool all motor outcomes of the upper limb, irrespective of whether they were measured with the FMA or the Chedoke-McMaster Stroke Assessment Scale. We assumed that this was appropriate because both scales are based on the same motor stages of Brunnstrom. Finally, only studies written in the English, German, or Dutch language were included. We may have inadvertently missed relevant studies that were not published in scientific journals or in other languages.

In summary, the present systematic review confirms the potential for robotic assisted devices to elicit improvements
in proximal upper limb function. However, improvements in terms of ADLs could not be substantiated. Unfortunately, administered ADL scales do not reflect recovery of the paretic upper limb properly. Future research on the effects of robot-assisted therapy should focus on kinematic analysis to differentiate between recovery by neural repair and recovery based on compensation strategies. In addition, trials should use valid instruments that measure upper limb skills specifically, such as Action Research Arm Test (ARAT) or Wolf Motor Function Test (WMFT). Finally, the cost-effectiveness of robotics must be investigated. This is particularly important because of the increasing pressure experienced by health care professionals in most countries to reduce health care costs. In addition, robotics may offer stroke patients an opportunity to train independently in an intensive functional fashion and at home. This becomes increasingly important when a lack of time is reported to be the major barrier for therapists not to comply with evidence-based guidelines for stroke rehabilitation.\textsuperscript{52,60}

**ACKNOWLEDGMENTS**

We thank students Ian Hoogendoorn, Vincent Groen, and Mario Kramer of the Faculty of Human Movement Science of VU University of Amsterdam for their participation in this research synthesis in collecting and reviewing papers. Dr. Krebs is supported by the NYS DOH Center for Research Excellence in Spinal Cord Injury and the US Public Health Service (NIH R01-HD045343, Veterans Affairs Baltimore B3688R and B3607R). Dr H.I. Krebs is a co-inventor in the MIT-held patents for the robotic devices. He holds equity positions in Interactive Motion Technologies, Inc., the company that manufactures this type of technology under license to MIT.

**REFERENCES**


* Items 2 to 11 of the PEDro scale.