

University of Groningen

Effectiveness of task-specific training using assistive devices and task-specific usual care on upper limb performance after stroke

Rozevink, Samantha G; Hijmans, Juha M; Horstink, Koen A; van der Sluis, Corry K

Published in:
Disability and rehabilitation. Assistive technology

DOI:
[10.1080/17483107.2021.2001061](https://doi.org/10.1080/17483107.2021.2001061)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Rozevink, S. G., Hijmans, J. M., Horstink, K. A., & van der Sluis, C. K. (2023). Effectiveness of task-specific training using assistive devices and task-specific usual care on upper limb performance after stroke: a systematic review and meta-analysis. *Disability and rehabilitation. Assistive technology*, 18(7), 1245-1258. <https://doi.org/10.1080/17483107.2021.2001061>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Effectiveness of task-specific training using assistive devices and task-specific usual care on upper limb performance after stroke: a systematic review and meta-analysis

Samantha G. Rozevink, Juha M. Hijmans, Koen A. Horstink & Corry K. van der Sluis

To cite this article: Samantha G. Rozevink, Juha M. Hijmans, Koen A. Horstink & Corry K. van der Sluis (2023) Effectiveness of task-specific training using assistive devices and task-specific usual care on upper limb performance after stroke: a systematic review and meta-analysis, *Disability and Rehabilitation: Assistive Technology*, 18:7, 1245-1258, DOI: [10.1080/17483107.2021.2001061](https://doi.org/10.1080/17483107.2021.2001061)

To link to this article: <https://doi.org/10.1080/17483107.2021.2001061>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 17 Nov 2021.



[Submit your article to this journal](#)



Article views: 7173



[View related articles](#)



[View Crossmark data](#)



Citing articles: 4 [View citing articles](#)

Effectiveness of task-specific training using assistive devices and task-specific usual care on upper limb performance after stroke: a systematic review and meta-analysis

Samantha G. Rozevink^a , Juha M. Hijmans^a , Koen A. Horstink^b and Corry K. van der Sluis^a 

^aUniversity of Groningen, University Medical Center Groningen, Department of Rehabilitation Medicine, Groningen, The Netherlands;

^bUniversity of Groningen, University Medical Center Groningen, Department of Human Movement Sciences, Groningen, The Netherlands

ABSTRACT

Purpose: Task-specific rehabilitation is a key indicator for successful rehabilitation to improve the upper limb performance after stroke. Assistive robotic and non-robotic devices are emerging to provide rehabilitation therapy; however, the effectiveness of task-specific training programs using assistive training devices compared with task-specific usual care training has not been summarized yet. Therefore, the effectiveness of task-specific training using assistive arm devices (TST-AAD) compared with task-specific usual care (TSUC) on the upper limb performance of patients with a stroke was investigated. To assess task specificity, a set of criteria was proposed: participation, program, relevant, repeated, randomized, reconstruction and reinforced.

Materials and methods: Out of 855 articles, 17 fulfilled the selection criteria. A meta-analysis was performed on the Fugl-Meyer Assessment scores in the subacute and chronic stages after stroke and during follow-up.

Results and conclusion: Both TST-AAD and TSUC improved the upper limb performance after stroke. In the sub-acute phase after stroke, TST-AAD was more effective than TSUC in reducing the upper limb impairment, although findings were based on only three studies. In the chronic phase, TST-AAD and TSUC showed similar effectiveness. No differences between the two types of training were found at the follow-up measurements. Future studies should describe training, device usage and criteria of task specificity in a standardized way to ease comparison.

ARTICLE HISTORY

Received 8 March 2021
Revised 3 September 2021
Accepted 26 October 2021

KEYWORDS

Stroke; upper limb; rehabilitation; task specific; assistive device

► IMPLICATIONS FOR REHABILITATION


- Arm or hand function is often undertreated in stroke patients, assistive training devices may be able to improve the upper limb performance.
- Task-specific training using assistive devices is effective in improving the upper limb performance after stroke.
- Task-specific training using assistive devices seems to be more effective in reducing impairment compared with task specific usual care in the subacute phase after stroke, but they are equally effective in the chronic phase of stroke.

Introduction

After stroke onset, the most reported deficiencies are related to motor dysfunction; 69% of the stroke patients admitted to the hospital experience upper limb deficits, of whom 32% have a severe paresis [1,2]. Engaging in upper limb rehabilitation can restore some of the arm function. Previous work has provided evidence that upper limb rehabilitation for stroke patients is most effective if it consists of repetitive, task-specific movements [3,4]. Task-specific training is characterized by training tasks that are meaningful to the patient, i.e., tasks that the patient would like to relearn [5]. Task-specific training is different from impairment-based training. Where task-specific training is goal directed and focussed on meaningful activities, impairment-based training aims to improve the impairments such as muscle weakness or

decreased range of motion [6]. Task-specific training has been found to be superior compared with impairment based training for functional improvement [7]. Unfortunately, a clear definition of task-specific training has not been adopted in the literature yet [4,8,9]. Hubbard et al. described task-specific training as: “training or intervention which utilizes ... ordinary everyday activities which are intrinsically and/or extrinsically meaningful to the patient or client” [4]. Hubbard et al. proposed five criteria to describe task-specific training, using the 5-R criteria; (1) training should be Relevant to the patient, (2) should be Repeated, (3) provided in a Randomized order, (4) contain tasks that are part of a whole task (Reconstruction) and (5) feedback should be provided (Reinforcement) [4]. The question rises on how repetitive and intensive task specific training can be provided in the most efficient way. Training using assistive robotic and non-robotic devices

CONTACT Samantha G. Rozevink  s.g.rozevink@umcg.nl  PO Box 30001, Groningen, 9700 RB, The Netherlands

 Supplemental data for this article can be accessed [here](#).

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

seems to be effective in improving the upper limb function of stroke patients due to the increased training intensity and duration and the option for home-use, as has been concluded in several systematic reviews [10–13]. These reviews showed that improvements were found on both the activity and the body functions and structure level of WHO's International Classification of Functioning, Disability and Health (ICF) model, meaning that upper limb performance improved [14]. When usual (non-robot assisted) care is provided in the same intensity as robot-assisted training, no significant differences in upper limb outcomes have been revealed yet [11]. However, the differences in device types and in training protocols make it hard to conclude whether device-assisted training is more effective in improving upper limb function compared with usual care. Interestingly, the previously mentioned reviews did not pay attention to the task specificity of the training protocols.

To determine how task specificity is implemented in the training programs of assistive robotic and non-robotic devices, we first need to look into the definition of what an assistive device is. We will use the term “assistive arm training devices (AATDs)” as an umbrella term to address robotic and non-robotic training devices. Similar to task specificity, there is a lack of a unified definition regarding AATDs. AATDs are only used for arm/hand training and not for support in daily life. Furthermore, AATDs can provide feedback during the training, can allow for unsupervised training and can be adjusted to match the training level of the patient [10,15]. Variables such as force or movement can be kept constant, which is impossible with human therapists. Many AATDs provide some type of anti-gravity support (to support the weight of the arm) or provide additional assistance to certain joints to perform a movement. In conclusion, we define AATDs as devices that support or assist (a part of) the upper limb to carry out a motion that is performed as (part of) a training program to improve upper limb performance.

The variety in devices has increased tremendously in the last two decades. AATDs can be roughly divided in end-effectors or exoskeletons [15]. End-effector devices are often only attached to the distal side of the body, allowing for instance the shoulder and elbow to move freely. Exoskeletons are aligned to the joints of the arm. AATDs can be used to play games or to perform other educational tasks to improve the upper limb function. Another feature that may differ between AATDs is the targeted body part. AATDs may focus on the proximal or distal parts of the upper limb or on the entire arm [16]. Lastly, AATDs can be roughly divided based on training modalities: passive, active-assistive, assistive, or active [16]. Note that Basteris et al. defined these modalities based on the level of patient participation during the training. In the passive modality, the patient's upper limb movement is entirely performed by the AATD. The active-assistive modality allows the AATD to sense that the patient is unable to complete the movement and will move the arm of the patient to the end goal. In assistive mode, the weight of the arm is supported. During active support, the device is not performing any movement but the patient is actively moving. Sensor based systems that are used solely for measurement purposes are active modalities.

According to Schweighofer et al., three principles should be applied when designing AATDs for task-specific training: promotion of skill acquisition of functional tasks, active participation of the patient and individualized training programs [8]. As mentioned before, Hubbard proposed the 5-R's: relevant, repeated, randomized, reconstruction and reinforced. To provide a definition of task specific training using an AATD, we combined the criteria

mentioned by Hubbard et al. and Schweighofer et al. We then noticed that the promotion of skill acquisition of functional tasks is also reflected in the criteria of Hubbard et al., especially in the relevance and reconstruction criteria. Therefore, this criterion provided by Schweighofer et al. was not included in our definition. Consequently, Schweighofer's “active participation” and “individualised program” criteria and Hubbard's criteria were combined to form a definition of task-specific training consisting of seven criteria: participation, program, relevant, repeated, randomized, reconstruction and reinforced. Hereafter called the PP5R criteria for task specificity assessment of training. Task specificity can be seen as a sliding scale; if more or less criteria are implemented in the training, the task specificity of the training will increase or decrease.

A fictional example of an AATD that would fulfil these criteria is a device that assists the patient while actively performing (participation) a relevant activity such as reaching and grasping of a virtual cup of tea to drink (relevant). The task can be broken down into separate components such as reaching and grasping, but in the end these components are combined to complete the entire task (reconstruction). The training is adapted to the progression of the patient in the amount of assistance provided (program). The position of the mug can be presented in a randomized order (randomized), and the exercises have to be practiced repeatedly (repetition) while receiving visual feedback on the performance (reinforced).

Often, the arm function declines over time when the AATD therapy is ceased. Evidence suggests that long-term outcomes from training with AATDs are not significantly different from conventional therapy [11]. Although task-specific training is better retained compared with non-task specific training [4], it has not yet been investigated if task-specific training using assistive arm devices (TST-AAD) is well retained on the long term.

The first aim of this systematic review was to investigate whether TST-AAD improved upper limb performance in stroke patients. The second aim was to compare TST-AAD to task-specific usual care (TSUC) training using meta-analysis. Finally, our third aim was to investigate how well the training of the upper limb performance was retained and if there was a difference in retention between TST-AAD and TSUC.

Materials and methods

This review was registered in Open Science Framework (OSF 10.17605/OSF.IO/QRA94) and the PRISMA guidelines were followed for reporting this review.

Eligibility criteria

The inclusion criteria were: (1) studies should concern adult humans diagnosed with stroke who as a result suffer from hemiplegia; (2) design: pilot, uncontrolled trial, crossover or randomized controlled trial (RCT) where the experimental group trained in a task-specific way (minimal 4 of the 7 of the PP5R criteria) using an assistive device and in case of a control group, this group trained in a task-specific way, where usual care was applied without using an assistive device; (3) upper limb exercises in the experimental group were performed using an AATD with either anti-gravity support or support of upper limb joint movement; (4) one of the outcome measures was an arm function test to quantify arm performance on ICF body function and structure or activity level and data were presented; (5) written in English or Dutch; (6) minimal number of participants in the study was 10 (sum of

experimental and control group); (7) primary research and published as full paper. Exclusion criteria were: (1) constraint-induced movement therapy (CIMT); (2) studies reporting on virtual environments without an assistive training device (like Kinect, or unadapted Wii), games, transcranial direct-current stimulation, brain-computer interface and functional electrical stimulation.

Literature search

The following databases were systematically searched: CINAHL, Embase, Cochrane and Pubmed. The search strategy was constructed in collaboration with a librarian of the Central Medical Library of the University Medical Centre Groningen. The following terms and synonyms were used: stroke, upper limb, robot, device, assistive technology, task specific. The search was limited to 31 October 2020, and there was no constraint on the start date. Detailed information concerning the search strategy per database can be found in Online [Supplementary Appendix I](#). One reviewer (SGR) performed the database search and removed duplicates. Two reviewers (CKvdS and SGR) performed the selection of articles based on title. If at least one of the reviewers included the paper during the title selection phase, the article was passed to the abstract phase. The title should at least mention the following words or synonyms: upper limb, stroke, training, device. Out of the remaining articles, the selection of abstracts was performed by two reviewers (SGR and KAH). Disagreements between the reviewers were discussed and if needed solved by a third reviewer (CKvdS). The selection of full text was performed by two reviewers (SGR and KAH), and disagreements were discussed with a third reviewer (CKvdS). The references of the included articles were screened for eligibility and if necessary added to the list. During the full text assessment, it was determined if the training programs provided to the experimental and control group were task specific. The PP5R criteria were applied [4,8]. The patient had to actively participate in the training, movements had to be initiated or completed (at least partly) voluntary. The program had to be individualized, meaning that if the patient improved, the task or exercise also had to become more challenging to maintain the same training intensity. A task was considered relevant to the patient if it had a clear link to activities of daily living (such as practising to pick up a cup or simulation of cleaning a table top) or had a clear context (in the case of gamification, the game had to have an engaging goal in an enriched environment). Therapy was randomized if at least the games or objects in the games were presented in a random order to the patient. Therapy was repetitive if the same movement or task was practiced subsequently more than twice. Reconstruction means that components of a task were practiced to work towards a more difficult task. In the case of “reaching for a cup”, reaching and grasping can be practiced separately and can eventually be combined. Lastly, reinforced means that some form of feedback on the movement performed was provided such as a score, visual feedback or other types of encouragement.

We introduced the PP5R criteria since, to our knowledge, there is no previous list of criteria to assess task specificity. Based on expert opinions, we adopted a threshold that at least four out of seven criteria should be described explicitly in the paper to be included in the review. Determining if a training was task specific appeared to be difficult, as is demonstrated by the following example. For instance “reaching” could be seen as a task on its own and could be seen as relevant if it was used as a component of a larger task, such as “reaching for a cup”. If so, it fulfils multiple items of the PP5R items of task-specific training. However,

reaching can also be seen as not relevant and not working towards reconstruction of a task if it is performed without context. In these cases, the training was regarded to be non-task specific. Examples of the latter were the training programs using the MIT-MANUS and its commercialized version, the InMotion, a robotic arm to train the shoulder and elbow, where participants had to reach for eight bulls-eye typed targets in a circular pattern [17–21]. This training program was not individualized, not randomized (the same movement was applied at all occasions: clockwise direction), not relevant in daily life and was not working towards a complete task since the task components remained the same.

Methodological quality

Methodological quality was assessed using the “Tool for the assessment of Study quality and reporting in EXercise” (TESTEX) scale [22]. This quality assessment scale is specifically developed for exercise intervention studies. It is partly based on the PEDro scale [23]. The PEDro scale is mostly used for exercise studies, but some items are redundant for these study types, such as the blinding of the participants which is often not possible in an exercise study. Other methodological factors, such as exercise intervention characteristics or the number of withdrawals, are important to report in an intervention study, but these factors are not part of the PEDro scale.

Methodological quality was assessed by two reviewers (SGR and KAH). Disagreements between the reviewers were discussed and if needed solved by a third reviewer (JMH). Three interpretations of the authors were added to the TESTEX evaluation: (1) the first item of the TESTEX concerns the eligibility criteria, for which is stated that all diagnostic values which are used to determine if a patient is included in the study must be reported and fulfilled. We defined that the diagnostic values for the inclusion criteria must be reported and fulfilled, but that the values for exclusion criteria were not required; (2) in item 10 of the TESTEX, it is stated that the activity of the sedentary control group must be monitored. The majority of the included studies did not have a sedentary control group since many patients received at least some therapy after stroke. Therefore, the item was interpreted as follows: it should be reported what type of exercises (frequency, duration, type of training) the control group executed during the intervention period; (3) a cut-off point to distinguish good from bad quality papers is not provided for the TESTEX [22]. Furthermore, the maximum score differs due to different designs. We therefore determined that studies with a score higher than 50% of their maximum score were of “sufficient quality” and considered for data analysis. Studies that scored lower than 50% of their maximum score were of “insufficient quality” and therefore excluded from data analysis. The cut off of 50% is similar to the cut off of the PEDro scale [24].

Meta-analysis

Review Manager (version 5.3) was used to calculate the standardized mean difference (SMD) for outcomes of the included studies [25]. A random effects model was used since the studies differed in intervention duration and device. Studies investigating an experimental group that trained in a task specific way using an AATD and a control group that trained in a task specific way without using an AATD were included. A meta-analysis was performed if three or more articles reported the same arm function test using a similar study design and if the data were presented in an

appropriate format (pre-post score or change to baseline and SD). Regarding the ICF body functions and structure level, the most common arm measurement was the Fugl-Meyer Assessment-Upper Extremity (FMA-UE). The FMA-UE is a performance test to determine the limitations in movement. For stroke studies, it is common to only report the motor function part of the FMA-UE which measures the upper arm (including the shoulder, elbow and lower arm), wrist, hand and speed/coordination (maximum score = 66) [26,27]. The minimal clinically important difference (MCID, a measure for a clinically relevant difference to the patient) was used to assess if improvement was meaningful. The MCID for the FMA-UE is six to eight points for patients in the chronic phase of stroke and nine points for patients in the sub-acute phase of stroke [28,29]. Regarding the ICF activity level, the Action Research Arm Test (ARAT) and the Wolf Motor Function Test (WMFT) were most commonly used. The ARAT investigates the arm and hand dexterity by picking up different items [26,30]. The maximum score is 57 points. The WMFT is comparable to the ARAT, it also measures the impairment and disability by interacting with different objects [31,32]. The WMFT has a Functional Ability Scale (FAS maximum score = 75) and a time scale (maximum score per item = 120). On the ICF participation level and activity level, the motor activity log (MAL) was most frequently applied. The MAL is a semi-structured interview where the patient is asked how often the arm is used in activities of daily living (ADL) tasks and how well the arm performed during executing

these tasks [33]. The mean score varies between 0 (never uses the arm) and 5 (use is same as before the stroke). In all aforementioned tests, higher scores indicate better performance of the upper limb, except for the time scale of the WMFT where a lower time represents faster execution of the task.

The mean and standard deviation of the change from baseline were used in the meta-analysis. When such results were not reported in an article, they were calculated [34]. A sensitivity analysis was performed by entering different values of correlation coefficients to investigate the effect on the standardized mean difference (SMD).

Results

Fifty-five articles were eligible after the screening procedure (Figure 1). Cohen's kappa was calculated to determine the agreement between the reviewers. Cohen's kappa of the two assessors who performed the full text assessment was 0.45, which means that their agreement was "moderate". This was mainly due to differences in interpretation of whether a training was task specific or not.

Task specificity

After the full text review, 40 articles were discarded, main reasons were non-task specific training or the task-specific training was

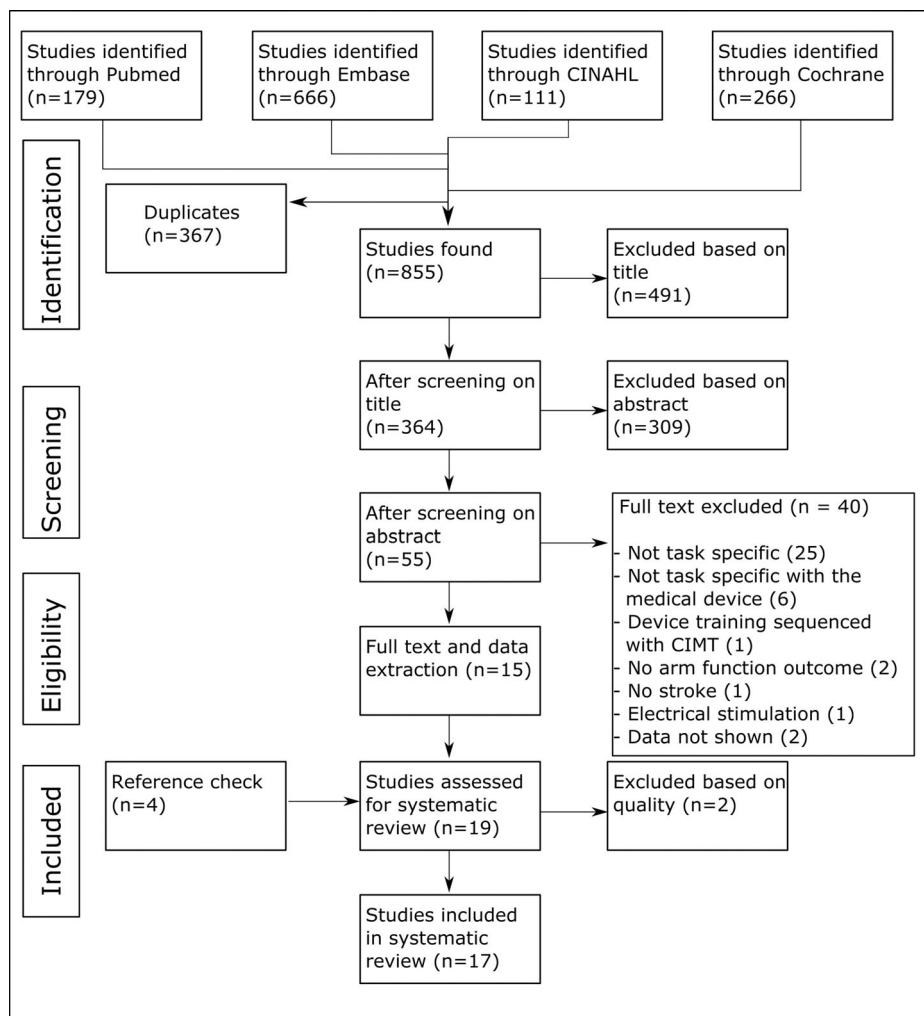


Figure 1. Flow diagram of the selection process according to the PRISMA guidelines.

Table 1. The assessment of task specificity of the training programs using the PP5R criteria in the included studies.

Author	Participation	Personalized program	Relevant	Random	Repetitive	Recon-struction	Reinforced	Sum
Bartolo et al. [36]	+	–	+	+	+	–	–	4/7
Byl et al. [37]	+	–	+	+	+	–	+	5/7
Casadio et al. [38]	+	+	–	+	+	+	+	6/7
Colombo et al. [39]	+	+	–	–	+	+	+	5/7
Connelly et al. [40]	+	+	+	+	+	–	+	6/7
Daunoraviciene et al. [41]	+	+	+	–	+	–	+	5/7
Fischer et al. [43]	+	+	+	+	+	+	+	7/7
Fluet et al. [44]	+	+	+	–	+	–	+	5/7
Housman et al. [45]	+	+	+	–	+	+	+	6/7
Huang et al. [46]	+	–	+	–	–	+	+	4/7
Klamroth-Marganska et al. [47]	+	–	+	–	–	+	+	4/7
Krebs et al. [35]	+	–	+	–	+	+	+	5/7
Lamberg et al. [48]	+	+	+	–	+	–	+	5/7
Page et al. [49]	+	+	+	–	+	+	+	6/7
Page et al. [50]	+	+	+	–	+	+	+	6/7
Takahashi et al. [51]	+	+	+	–	+	–	+	5/7
Thielbar et al. [52]	+	+	+	–	+	+	+	6/7
Timmermans et al. [39]	+	+	+	+	+	+	+	7/7
Tomić et al. [42]	+	+	+	+	+	+	+	7/7
Sum	19/19	14/19	17/19	7/19	17/19	12/19	18/19	

Note: Task specificity is based on the PP5R criteria which consist of the criteria mentioned in Hubbard et al. [4] and Schweighofer et al. [8]. Sum scores are provided within studies and across studies. + met the criteria; –not met the criteria.

not performed with an AATD. Most studies using the MIT-MANUS were excluded during the task-specific assessment; however, the study of Krebs et al. [35] was included since they used the MIT-MANUS in a task-specific way by using ADL objects for virtual and physical grasping movements.

For 19 studies, the task-specific criteria were assessed. Three studies scored all seven PP5R criteria (Table 1). Six studies scored six points, seven studies scored five points and three studies scored four points. The least reported item was randomization of the training. In many gaming environments, the exercises would probably be presented in a random order/way, but only four studies reported that items were randomly presented in the gaming environment [36,38,40,53]. Two included studies had multiple experimental groups, and these groups were only different in amount of support that was provided by the device [37,51]. The type of training was similar, therefore no distinction was made in the task specificity scoring of these different exercise groups.

Quality assessment

Two reviewers performed the TESTEX assessment (Table 2). Their agreement was reflected in a Cohen's Kappa of 0.75 indicating a "substantial" agreement. Two studies were of insufficient quality and were therefore disregarded from further analysis [35,39]. Four studies did not include a control group or the experimental group was divided based on time after stroke and therefore could not report a randomization procedure [38,42–44]. If studies did not have any drop outs, an intention to treat analysis (ITT) was not applicable and this TESTEX item was not assessed. If randomization was not applicable or there were no dropouts, the maximum number of points that could be assigned was reduced. None of the studies scored all points. The least reported item was the ITT, but many studies did not have any drop outs so this item was not applicable in most cases. The most reported item was the exercise adherence, meaning that patient drop out was less than 15%.

Characteristics of the included studies

Relevant data were extracted from the remaining 17 studies, in which in total 383 participants were involved (Table 3). Except for

four studies reporting on the sub-acute phase after stroke ($n=101$) [36,41–43], participants were in the chronic phase of stroke (>6 months post stroke, $n=282$). Fifteen different robotic devices were used. In 11 studies, training was provided via a computer screen or virtual reality, in six studies training used real-life objects. Thirteen studies reported on one or more experimental groups and a control group, four studies reported on one or more experimental groups without a control group. The total number of participants in the experimental groups was 221, and 162 participants were assigned to the control groups.

The mean (\pm SD) age was 58.2 ± 10.3 years and 57.2 ± 7.9 years in the experimental and control groups, respectively. The mean time post stroke was 3.7 ± 1 years and 4.5 ± 4.4 years in the experimental and control groups, respectively (not reported in [36,50]). Slightly more often a hemiparesis on the left side (191) was observed compared with the right side (192). More men (226) than women (157) participated in the research. The total amount of training time varied from 6 to 32 h.

Outcomes of the training programs

In all studies, the arm performance improved in the experimental group compared with baseline in one or multiple tests (Table 4). A significant improvement in arm performance was also observed in 11 of the 13 control groups.

In 3 of the 13 included RCTs, the experimental group improved significantly more than the control group in one or more functional outcomes [42,47,49]. Improvements could not be attributed to a specific robotic device or the number of training hours. In five studies, the improvement in arm function on one or more outcomes in the control group was larger, although not significant, than the experimental group [36,47,50,52,53].

One experimental group improved significantly more compared with another experimental group [51]. The study of Takahashi et al. showed that a group that trained only in the active-assistive mode gained more improvement in arm performance than the group that trained half of the time in active-assistive mode and half of the time in active non-assistive mode [51]. Three pre-posttest studies were included in which the patients significantly improved the arm performance after training [38,43,48].

Table 2. Qualitative assessment of the included studies using TESTEX.

Author	Study quality (max 5)					Study reporting (max 10)							Total		
	Eligibility (1)	Randomization (1)	Allocation Concealment (1)	Groups equal (1)	Blind assessor (1)	Exercise adherence (85%) (1)	Adverse event (1)	Session attendance (1)	ITT analysis (1)	Between-group comparisons (2)	P value and variability (1)	Control group activity (1)		Constant intensity (1)	Exercise volume and expenditure (1)
Bartolo et al. [36]	-	+	-	+	+	+	-	-	nd	++	-	+	-	+	8/14
Byl et al. [37]	-	-	+	+	+	-	+	+	-	+-	+	-	-	+	8/15
Casadio et al. [38]	+	na	na	na	-	+	-	+	-	na	+	na	+	+	6/9
Colombo et al. [39]	-	na	-	-	-	+	-	-	nd	--	+	na	+	-	3/12
Connelly et al. [40]	+	-	-	+	-	+	+	-	nd	+-	+	+	+	+	9/15
Dauroraviciene et al. [41]	-	-	-	+	-	+	-	+	nd	++	+	-	+	-	7/14
Fischer et al. [43]	+	na	na	na	-	+	-	-	-	++	+	-	+	+	8/12
Fluet et al. [44]	+	-	-	+	-	+	+	+	nd	++	+	+	+	+	11/14
Housman et al. [45]	+	+	+	-	+	+	+	+	-	++	+	+	+	+	13/15
Huang et al. [46]	-	-	+	+	-	+	-	+	nd	++	+	+	-	+	9/14
Klamroth-Marganska et al. [47]	+	+	+	+	+	+	+	+	+	++	+	+	-	+	14/15
Lambercy et al. [48]	-	na	na	na	+	+	+	+	-	na	+	na	+	+	7/9
Krebs et al. [35]	-	-	-	-	+	+	-	-	nd	+-	+	+	-	+	6/14
Page et al. [49]	-	+	+	-	+	+	-	-	nd	++	+	+	+	+	11/14
Page et al. [50]	-	+	+	-	+	+	-	-	nd	++	+	+	+	+	10/15
Takahashi et al. [51]	-	-	-	+	-	+	-	-	nd	++	+	+	+	+	9/14
Thielbar et al. [52]	+	+	+	+	+	+	-	-	nd	++	+	+	+	+	11/14
Timmermans et al. [53]	-	+	+	+	+	+	+	+	nd	++	+	+	+	+	13/14
Tomic et al. [42]	+	+	-	+	+	+	+	-	nd	++	+	+	+	+	12/14
Total:	8	8	8	11	11	18	11	9	1	13	17	13	14	17	

Note: + met the criteria; -not met the criteria.

na: not applicable; nd: no dropout; ITT: intention to treat analysis; TESTEX: tool for the assessment of Study quality and reporting in Exercise.

Meta-analysis

FMA-UE

Sixteen studies reported the FMA-UE, but four studies had a pre-posttest design and could not be included in the meta-analysis [38,43,48,51]. Two studies were discarded from the meta-analysis because they either used medians and interquartile ranges or an adapted version of the FMA-UE [44,53]. The meta-analysis of the 10 remaining studies [36,37,40–42,45,47,49,50,52] was divided in a sub analysis of studies addressing the subacute phase and chronic phase of stroke separately (Figure 2). The meta-analysis for the subacute phase revealed that the upper limb impairment in the experimental groups reduced significantly more than the control groups, despite the fact that in the three studies the duration of training was limited to 7.5 h in 3 weeks, 10 h in 2 weeks and 6 h in 2 weeks, respectively [36,41,42]. Kwakkel et al. advise that at least 16 h of training is necessary for improvement [54]. In the chronic phase of stroke, decrease in upper limb impairment was not significantly different between the experimental groups and control groups.

A meta-analysis was not possible for the WMFT, ARAT and MAL. Although multiple papers included these outcomes (Table 4), the data were reported in different ways and could not be used to calculate the change from baseline and the standard deviation.

Meta-analysis on follow-up outcomes FMA-UE

Only the FMA-UE was used for the meta-analysis on the follow-up results, since this was the only outcome measure with sufficient data. Nine studies included a follow-up measurement including the FMA-UE, varying in time after the intervention was finished (range 4 weeks to 36 weeks). The design and data of four studies were appropriate to perform a meta-analysis [40,45,47,52]. The follow-up measurements in the included studies took place between 4 and 34 weeks after cessation of the intervention. The study of Klamroth-Marganska et al. was included twice in the meta-analysis, since these authors performed a follow-up after 16 and 34 weeks [47]. The meta-analysis showed that the experimental

group did not perform significantly better than the control group on the FMA-UE (Figure 3).

Discussion

Applying TST-AAD can lead to improvement of the upper limb performance in stroke patients, since all studies showed an improvement in one or more motor function or impairment tests, regardless of the phase after stroke (subacute or chronic). TST-AAD seems to have added value in the subacute phase since larger improvements were observed compared with the control groups. In the chronic phase, TST-AAD had no additional effect over TSUC. No significant differences between the TST-AAD and TSUC treatments could be revealed in the follow-up phase.

Our primary aim was to determine if TST-AAD of stroke patients was effective in improving the upper limb performance. We showed that the application of TST-AAD can improve the upper limb performance in stroke patients. To be able to proof effectiveness of TST-AAD, it is important to know the criteria for task specificity. Although Hubbard et al. and Schweighofer et al. defined criteria for “task specific training”, it appeared that these are not yet widely used in the literature [4,8]. In many papers it was stated that task-specific training was provided; however, a detailed explanation of what the therapy included was lacking. To assess these criteria retrospectively appeared to be very difficult, whereas it is easy to prospectively include the PP5R criteria in future manuscripts. Currently, the content of training protocols described in literature is in general too concise and lack relevant information. The training protocol in the methods section should include the criteria of task specificity so the reader is able to assess if the training indeed fulfilled the criteria of task-specific training. Furthermore, the number of repetitions is important to include since this is an important indicator for determining the intensity of the training. The number of repetitions during one complete training session was only reported in five of the included studies.

Task-specific training seemed to be effective to improve the upper limb function and reduce impairment, in all studies at least one of the outcomes showed significant improvements after

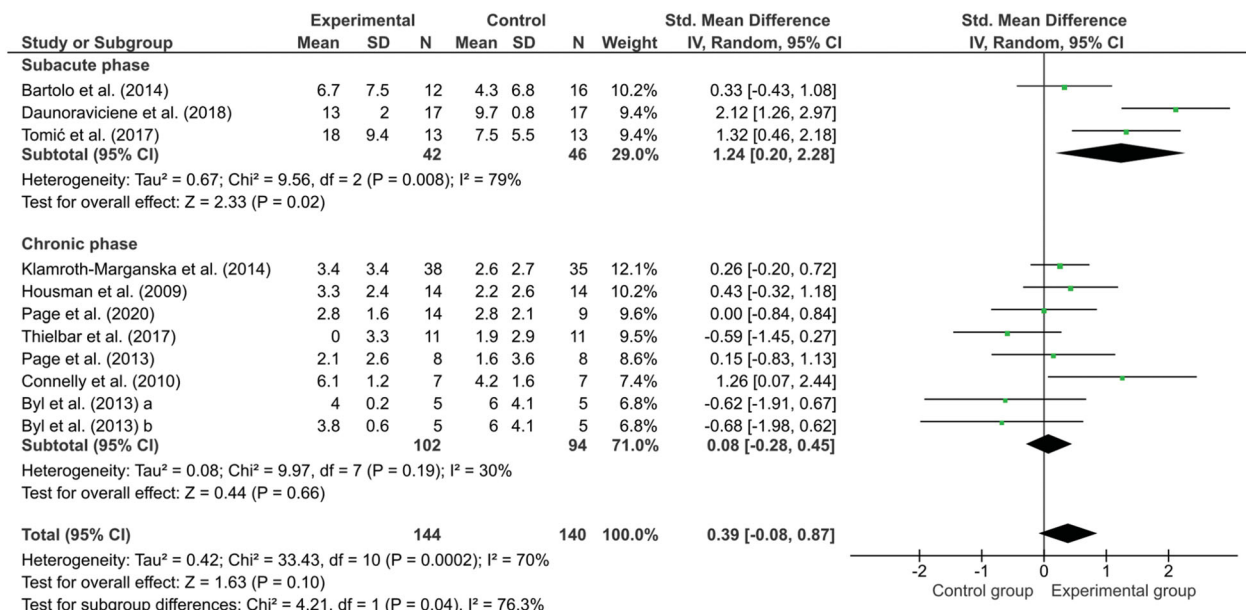


Figure 2. Meta-analysis of Fugl-Meyer Assessment scores. Change from baseline was compared between the experimental group (task-specific assistive devices) and the control group (task-specific usual care training without device). Byl et al. contained two exercise groups (a = unilateral; b = bilateral).

Table 3. The characteristics of the studies included in the review.

Authors	Design	Group	Assistive device + equipment	Exoskeleton or end-effector	Training modality [#]	Targeted joints	N (M/F)	Age (mean years ± sd)	Affected side (L/R)	Time since stroke (mean years ± sd)	Repetitions per session	Training hours
Bartolo et al. [36]	RCT	EG CG	Armeo Spring + screen intensive TSUC	Exoskeleton	Assistive	SH/EL/WR	12 (8/4) 16 (10/6)	63.5 ± 11.6 51.2 ± 7.0	6/6 7/9	nr nr	nr nr	6 18
Byl et al. [37]	RCT	EG1 EG2	UL-EXO7 unilateral + screen UL-EXO7 bilateral + screen	Exoskeleton	Assistive, active-assistive and passive mirrored (dependent on game)	SH/EL/WR	5 (4/1) 5 (5/0)	54.2 ± 20.5 65.2 ± 5.4	2/3 1/4	10.2 ± 5.0 8.4 ± 4.2	nr nr	18
Casadio et al. [38]	Pre-post	CG	intensive TSUC	End-effector	Assistive	SH/EL	5 (4/1)	59.3 ± 6.8	2/3	6.4 ± 4.4	nr	10
Connelly et al. [40]	RCT	EG EG CG	Braccio di Ferro + screen PneuGlove + VR VR only with TST	Exoskeleton	Active-assistive	Fingers	10 (3/7) 7 (3/4) 7 (3/4)	52.8 ± 13.4 54 ± 10 57 ± 18	5/5 4/3 3/4	3.9 ± 3.5 4.8 ± 1.5 10.2 ± 11.8	63 to 126 97 119	18
Dauroravičienė et al. [41]	RCT	EG CG	Armeo Spring + screen intensive TSUC	Exoskeleton	Assistive	SH/EL/WR	17 (11/6) 17 (11/6)	65.9 ± 4.9 65.5 ± 4.1	6/11 4/13	8.6 ± 3.5 weeks* 9.7 ± 6.2 weeks*	nr nr	10
Fischer et al. [43]	Pre-post	EG	X-Glove + real world objects	Exoskeleton	Active-assistive	WR/Fingers	13 (10/3)	63 ± 12	6/7	4 ± 1 months*	163	15
Fluet et al. [44]	nRCT	EG CG	Haptic MASTER + screen TSUC	End-effector	Active-assistive	SH/EL/WR/Fingers	10 (6/4) 11 (4/7)	53 ± 11 50 ± 11	6/4 4/7	5.0 ± 3.4 7.3 ± 4.7	nr nr	24
Housman et al. [45]	RCT	EG CG	T-WREX + screen intensive TSUC	Exoskeleton	Assistive	SH/EL/Fingers	14 (11/3) 14 (7/7)	54.2 ± 11.9 56.4 ± 12.8	10/4 10/4	7.0 ± 8.0 9.4 ± 10.7	nr nr	24
Huang et al. [46]	RCT	EG	3D-DHD + real world objects	Exoskeleton	Assistive and resistance	Fingers	5 (5/0)	59.6 ± 8.0	3/2	4.3 ± 3.5	nr	14
Klamroth-Marganska et al. [47]	RCT	CG EG	intensive TSUC ARMin + screen	Exoskeleton	Assistive	SH/EL/Fingers	5 (4/1) 38 (17/21)	56.6 ± 7.9 55 ± 13	4/1 19/19	4.1 ± 2.2 4.5 ± 3.7	nr nr	18
Lamberg et al. [48]	Pre-post	CG	intensive TSUC	End-effector	Assistive or resistance	Fingers	15 (7/8)	58 ± 14	19/16	3.3 ± 3.8	nr	18
Page et al. [49]	RCT	EG CG	HapticKnob + screen Myomo e100 + real world objects	End-effector	Assistive	EL	8 (3/5) 8 (8/0)	55.5 ± 14.6 59.0 ± 12.9	11/4 4/4	1.6 ± 0.8 3.7 ± 3.2	nr nr	24
Page et al. [50]	RCT	EG CG	intensive TSUC Myomo + real world objects	End-effector	Assistive	EL	8 (8/0) 14 (9/5)	58.5 ± 9.5 55.8 ± 9.3	5/3 9/5	8.9 ± 9.6 nr	nr nr	24
Takahashi et al. [51]	Pre-post	EG1 EG2	HWARD assistive + screen HWARD 1/2 of the time assistive + screen	End-Effector	Active-assistive or active	WR/Fingers	9 (5/4) 7 (3/4) 6 (3/3)	57.2 ± 7.7 58.6 ± 16 67.3 ± 15	7/2 0/13	nr 1.2 ± 1.1 4.8 ± 7.3	nr only first half of training = 90	15
Thielbar et al. [52]	RCT	EG CG	VAEDA Glove + real world objects intensive TSUC	Exoskeleton	Assistive	Fingers	11 (7/4) 11 (8/3)	61 ± 12 56 ± 10	6/5 3/8	7.9 ± 9.5 4 ± 3.9	146 174	18
Timmermans et al. [53]	RCT	EG CG	Haptic Master + real world objects intensive TSUC	End-effector	Active-assistive	SH/EL/WR/Fingers	11 (8/3) 11 (8/3)	61.8 ± 6.8 56.8 ± 6.4	7/4 8/3	2.8 ± 2.9 3.7 ± 3.0	nr nr	32
Tomic et al. [42]	RCT	EG CG	ArmAssist + screen intensive TSUC	End-effector	Active	SH/EL	13 (12/1) 13 (9/4)	56.5 ± 7.4 58.3 ± 5.2	5/8 6/7	35.3 ± 9.7 days* 37.3 ± 7.7 days*	120-150 120-150	7.5

CG: control group; EG: experimental group; EL: elbow; L: left; N: number of participants (M: male/F: female); nr: not reported; (n)RCT: (non) randomized controlled trial; R: right; SD: standard deviation; SH: shoulder; TST: task-specific training; TSUC: task-specific usual care; VR: virtual reality; WR: wrist; #: Classification according to Basteris et al.; *: Patients in the subacute phase after stroke. 3D-DHD: 3D (printed) dynamic hand device; HWARD: hand wrist assistive rehabilitation device; MEMOS: mechatronic system for motor recovery after stroke; MIT: Massachusetts Institute of Technology; T-WREX: therapy Wilmington robotic exoskeleton; VAEDA: voice and EMG-driven actuated.

Table 4. The baseline and change to baseline measures of the included studies.

Author	Assistive device	Outcome measure	Baseline score (mean ± sd)		Change to baseline (mean ± sd); comparison EG versus CG	
			EG‡	CG	EG‡	CG
Bartolo et al. [36]	Arm weight support device	FMA-UE	22.6 ± 21.8	27.8 ± 18.9	6.7 ± 7.5 [#]	4.3 ± 6.8 [#]
		FIM	40.3 ± 10.0	43.1 ± 19.6	36.8 ± 18.1 ^{*#}	40.5 ± 13.8 ^{*#}
Byl et al. [37]	UL-EXO7	FMA-UE	23.4 ± 7.7 UL	24.6 ± 6.8	4.0 ± 0.2 ^{* UL}	6.0 ± 4.1 ^{*§}
			24.4 ± 5.2 BL		3.8 ± 0.6 ^{* BL}	
			WMFT, B&B, Finger tapper, DRTT: Not reported per group, only for all subjects combined (EG1 + EG2 + CG)			
Casadio et al. [38]	Braccio di Ferro	FMA-UE	15.0 ± 13.0		4.5 ± 2.6 [*]	
Connelly et al. [40]	PneuGlove	FMA-UE	37 ± 3.3	38.4 ± 4.5	6.1 ± 1.2 ^{*§}	4.2 ± 1.6 ^{*#}
		B&B	16.4 ± 2.1	20.6 ± 2.6	3.4 ± 0.8 ^{*#}	0.8 ± 0.9 ^{*#}
Daunoraviciene et al. [41]	Armeo Spring	FMA-UE	32.2 ± 16.5	32.1 ± 16.2	13.0 ± 2.0 ^{(*)§}	9.7 ± 0.8 ^{(*)§}
		FIM	24.4 ± 5.2	25.8 ± 8.2	7.53 ± 0.8 ^(*)	2.0 ± 0.5 ^(*)
Fischer et al. [43]	X-Glove	FMA-UE	17.0 ± 6.1		2.2 ± 2.1 ^{*#}	
		ARAT	10.5 ± 6.2		1.9 ± 3.0 ^{*#}	
		CAHAI-9	1.8 ± 0.8		0.6 ± 0.4 ^{*#}	
		WMFT-FAS	2.7 ± 1.1		0.4 ± 0.5 ^{*#}	
		WMFT-Time	68.3 ± 29.4		-12.1 ± 10.5 ^{*#§}	
		MAL QoM	0.6 ± 0.5		0.6 ± 0.4 ^{*#}	
Fluet et al. [44]	Haptic MASTER	FMA-UE ¹	48 ± 5.6	52 ± 8.7	2.3 ± 4.8	1.9 ± 2.7
		WMFT-Time ²	6.2 ± 1.8	4.6 ± 2.4	-1.6 ± 1.7 [*]	-0.9 ± 1.1 [*]
Housman et al. [45]	T-WREX	FMA-UE	21.7 ± 5.9	18.1 ± 5.0	3.3 ± 2.4 [*]	2.2 ± 2.6 [*]
		Rancho	3.4 ± 0.6	3.4 ± 1.0	0.1 ± 0.4	0.1 ± 0.7
		MAL AoU	0.6 ± 0.4	0.3 ± 0.3	0.2 ± 0.4	0.1 ± 0.3
		MAL QoM	0.6 ± 0.4	0.3 ± 0.3	0.2 ± 0.4	0.2 ± 0.3 [*]
Huang et al. [46]	3D-DHD	B&B	6.8 ± 7.6	8.6 ± 6.8	6.8 ± nr ^{*§}	2.2 ± nr [*]
Klamroth-Marganska et al. [47]	ARMin	FMA-UE	20.2 ± 7.1	20.7 ± 8.2	3.4 ± 3.4 ^{(*)†}	2.6 ± 2.7 ^(*)
		WMFT, MAL and SIS: Only reported F-ratio				
Lambercy et al. [48]	HapticKnob	FMA-UE	32.2 ± 11.3		3.0 ± 4.1 [*]	
		MAS	4.1 ± 3.3		1.0 ± 1.1 [*]	
		FTHUE	2.9 ± 0.6		0.2 ± 0.2	
Page et al. [49]	Myomo e100	FMA-UE	20.8 ± 4.8	19.4 ± 4.6	2.1 ± 3.0 [#]	1.6 ± 3.6 [#]
		SIS	216 ± 25.5	239.4 ± 17.5	23.9 ± 8.7 ^{(*)#†}	8 ± 8.4 ^{(*)#}
		COPM	4.5 ± 2.2	5.2 ± 2.4	-0.3 ± 0.8 [#]	-0.4 ± 1 [#]
Page et al. [50]	Myomo	FMA-UE	19.6 ± 3.8	17.6 ± 4.4	2.8 ± 1.6 ^{(*)#}	2.8 ± 2.1 ^{(*)#}
		AMAT	28.4 ± 7.3	22.8 ± 6.6	0.9 ± 3.2 [#]	2.6 ± 2.4 ^{(*)#}
Takahashi et al. [51]	HWARD	FMA-UE	40.4 ± 10.5 A		9.1 ± 2.1 ^{*§† A}	
			49.5 ± 8.6 HA		5.8 ± 1.6 ^{* HA}	
		ARAT	28 ± 20 A		5.3 ± 2.1 ^{*† A}	
			41 ± 19 HA		2.8 ± 1.8 ^{* HA}	
		B&B	15 ± 19 A		5.0 ± 4.2 ^{* A}	
			25 ± 19 HA		3.3 ± 2.4 ^{* HA}	
		NHP	47 ± 38 A		-9.6 ± 16.6 A	
			39 ± 22 HA		-3.0 ± 3.8 HA	
		SIS	1.9 ± 0.9 A		0.9 ± 0.4 ^{*† A}	
			3.1 ± 1.1 HA		0.3 ± 0.3 HA	
Thielbar et al. [52]	VAEDA Glove	FMA-UE	34.8 ± 9.7	26.5 ± 8.5	0 ± 3.3 [#]	1.9 ± 2.9 [#]
		ARAT	33.7 ± 13.5	26.6 ± 10.9	2.6 ± 4.8 ^{*#}	1.0 ± 4.5 [#]
		WMFT-Time	31.2 ± 26.0	35.2 ± 21.4	-10.1 ± 13.9 ^{*#}	-3.5 ± 7.7 [#]
		CMSA-H	4.0 ± 0	4.0 ± 0	-0.1 ± 0.3 ^{*#}	1 ± 0.3 [#]
Timmermans et al. [53]	Haptic Master	FMA-UE ³	53 ± [39–58]	50 ± [47–57]	5 [nr]	1 [nr]
		ARAT ³	31 ± [24–40]	39 ± [28–46]	3 [nr] [*]	4 [nr]
		MAL AoU ³	2.2 ± [1.5–2.8]	2.2 ± [1.5–3.8]	0 [nr]	1.5 [nr] [*]
		MAL QoM ³	1.5 ± [1.1–2.1]	1.8 ± [0.9–3]	0.1 [nr] [*]	1.2 [nr] [*]
Tomić et al. [42]	ArmAssist	FMA-UE	26.5 ± 7.7	26.6 ± 7.5	18.0 ± 9.4 ^{(*)§†}	7.5 ± 5.5 ^(*)
		WMFT-FAS	44.2 ± 12.2	42.4 ± 13.3	14.1 ± 7.9 ^{(*)§†}	6.7 ± 7.8 ^(*)
		BI	65.0 ± 26.1	65.4 ± 19.8	21.2 ± 24.8 ^{(*)§}	13.1 ± 10.7 ^{(*)§}

Baseline and change from baseline. A: assistive; AMAT: arm motor activity test; ARAT: action research arm test; B&B: box and block test; BI: Barthel index; BL: bilateral; CAHAI-9 = Chedoke arm and hand inventory; CG: control group; CMSA-H: Chedoke McMaster stroke assessment-hand; COPM: Canadian occupational performance measure; EG = exercise group; FAS: functional ability scale; FIM: functional independence measure; FMA-UE: Fugl-Meyer assessment-upper extremity; FTUUE: functional test of hemiparetic upper extremity; HA: half of the time assistive; IQR: interquartile range; MAL: motor activity log (AoU: amount of use; QoM: quality of movement); MAS: motor assessment scale; MSS: motor status scale (sh: shoulder; ha: hand); NHP: nine hole peg test; nr: not reported; PR: progressive-resistive; Rancho: Rancho Los amigos functional test of upper extremity function; sd: standard deviation; SIS: stroke impact scale; SM: sensorimotor; UL: unilateral; WMFT: Wolf motor function test. 3D-DHD: three-dimensional dynamic hand device; 1 = max score 64; 2 = mean time; 3 = median and interquartile range; *: Significantly different from baseline; (*): Significance not calculated in article and therefore a paired t-test performed; †: significantly different from control group; ‡: For multiple exercise groups, see Table 3; ¶: FMA-UE including sensation, passive joint motion and joint pain; #: SD calculated using coefficient correlation; §: clinically relevant improvement.

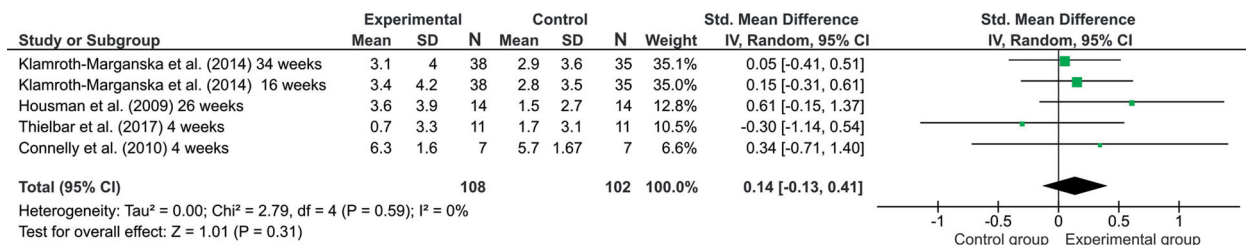


Figure 3. Meta-analysis of the follow-up results of the Fugl-Meyer Assessment. Study of Klamroth-Marganska et al. [47] performed two subsequent follow-up tests.

training. A significant improvement on the FMA-UE between baseline and post intervention in the experimental group was observed in 11 out of 16 studies that measured this outcome. The studies that did not find a significant improvement in FMA-UE were different in time after stroke, device and training, therefore any specific reason why there were no significant improvements in FMA-UE in these studies cannot be provided. Only in two studies the changes in FMA-UE were significantly higher after TST-AAD compared with TSUC [42,47]. The study population, type of device, training and duration of the training were different between these studies. The only similarity was the targeted body part; the shoulder and elbow. However, other studies targeting the shoulder/elbow did not reveal significant differences between the experimental and control groups. Therefore, it remains unclear why aforementioned studies were more successful.

Interestingly, the majority of TSUC control groups were also effective in improving the arm performance, which supports the idea that task-specific training improves upper limb performance, regardless if it is provided by a therapist or a device [9,55–57]. This finding is of clinical relevance, since assistive device training can be executed with limited therapist assistance or can be executed at home. When provided in combination with a telerehabilitation platform, patients have the opportunity to train at home more frequently and independently, which may save therapists' time. Further research should therefore focus on the cost-efficiency aspects of robotic versus individual task-specific therapy.

TST-AAD appeared to be more effective than TSUC in the subacute stage of stroke, but not in the chronic phase. However, the effectiveness was based on only three studies in the subacute phase. More RCTs in the subacute phase should therefore be performed to investigate the effect of TST-AAD on the upper limb in this stage. In the chronic phase of stroke, the effectiveness of TST-AAD was comparable to TSUC. The studies included in the meta-analysis reporting positive FMA-UE effects of TST-AAD over TSUC in the chronic phase included younger participants [40,45,47]. The brains of younger patients are more plastic compared with older aged brains, which could be a reason why the studies showed a positive effect of TST-AAD [58]. Furthermore, TST-AAD is more challenging, new and exciting compared with TSUC, creating an enriched environment from which it is known to increase neurogenesis. Neurogenesis decreases with age, which may explain why the studies with older patients did not find an effect of TST-AAD.

Although no significant differences between experimental and control groups could be demonstrated with regard to the retention effects, a trend in favour of the experimental group seemed apparent. Although most studies showed a larger improvement in the experimental group, differences were small and only in the study of Housman et al., this difference was statistically significant [45]. Whether this suggests that the TST-AAD seems to be better retained than TSUC should therefore be investigated further.

Our finding that robotic training is as effective as usual care is in accordance with other reviews [10–13]. Our review is different since we focussed specifically on task specificity, whereas other reviews included all types of robotic training. Task specificity was briefly mentioned in previous reviews, but was not used for study selection so far [10–13]. Norouzi-Gheidari et al. described in their review that robotic therapy was more effective if it was provided in addition to usual care [11]. They did not find a significant difference when robotic therapy and conventional therapy performed the same tasks in the same amount of time, as we have also shown in our results. In contrast, the review of Bertani et al. showed that robotic therapy may be more effective in improving the arm function, especially in chronic stroke [13]. The authors could not explain why the chronic stroke group improved significantly more than usual care, nor did they compare their outcome to other reviews. Differences in inclusion criteria of the studies may explain the conflicting conclusions.

Only a few studies demonstrated clinically meaningful improvements after applying TST-AAD according to known MCIDs. In two studies, the increase in FMA-UE score was larger than the MCID for patients in the subacute phase after stroke, although we should keep in mind that this improvement was also partly attributed to spontaneous recovery [41,42]. Interestingly, these two studies had the lowest training time of all included studies but the largest improvements in FMA-UE. Future studies with these devices could investigate longer intervention periods which may lead to even larger reductions in impairment. In the chronic phase of stroke spontaneous recovery is thought to have reached its limits, observed improvements can therefore be more certainly attributed to training [59–61]. The FMA-UE MCID for patients in the chronic phase after stroke is six to eight points [29]. Three studies showed FMA-UE improvements larger than the MCID [37,40,51]. An interesting observation was that in two studies, the patients whose improvements exceeded the MCID were mostly moderately mildly impaired as measured with the FMA-UE (baseline scores >37) [40,51,62]. This could indicate that patients with mild impairments in the chronic phase can gain more from TST-AAD than more severely affected patients. In previous studies, it has also been reported that mildly impaired chronic patients benefit more from training with AATDs [63,64]. However, a systematic review to support this hypothesis is lacking so far.

The variation in AATDs makes it challenging to compare the studies. Large differences in therapy effects were observed between devices that target the same function. For instance, the HWARD and VEADA Glove both target the fingers. Although training using the HWARD showed a clinically relevant increase on the FMA-UE, the VEADA Glove did not result in improvements [51,52]. The EMG-driven devices seemed to have the least effect on improving the arm function (Myomo and VEADA Glove), and this may be explained by the fact that EMG-driven devices can be less reliable due to movement artefacts or to different muscle activation after stroke [49,52,65,66]. In the chronic phase, the largest

improvements in upper limb function were observed with devices that trained the fingers (HWARD and PneuGlove) [40,51]. A possible explanation could be that the corticospinal tracts from the fingers to the cortex may be suppressed after stroke [67]. The part surrounding the infarction (penumbra) will heal and damaged corticospinal tracts in this part of the brain will become useful again. Also, other distant connecting unaffected parts of the brain will change due to the damage in the injured part of the brain (diaschisis). Since these phenomena take up several weeks or even months, the patient may be under the impression that the fingers were not useful, resulting in non-use [68]. In the chronic phase of stroke, the penumbra and diaschisis may be (partially) restored. When patients take part in a study that intensively trains the finger function, the patient may progress rapidly due to using the fingers again which could be due to a behavioural change or to plasticity of the brain. Plasticity of the brain is also possible in chronic stroke, as was shown by Szaflarski et al. and Liepert et al. [69,70]. By forcing to use the affected hand (during CIMT), the cortical representation of the affected hand was partially restored. Similar findings have been reported for robotic training, arguing that even in the chronic phase plasticity of the brain may be possible [51,71]. In contrast, in the subacute phase of stroke, the largest improvements in upper limb function were observed after application of the ArmAssist and the Armeo Spring, focussing on self-initiated movements of the shoulder and elbow using anti-gravity support [41,42]. In the subacute phase after stroke, training of proximal muscles that are used for gross movements improve earlier than the distal muscles, due to the bilateral innervation and non-corticospinal input that proximal muscles receive [60]. However, one should take into consideration that there is a continuing competition for representation on the motor cortex [71,72]. Training of solely the proximal parts of the upper limb could reduce the representation of the hand on the cortex, as has been argued by Hluštík and Mayer [71]. They advise to start training the hand as early on as possible in the course of rehabilitation. Combined training of the shoulder and hand might be the most stimulating way of training and should be further investigated in the subacute phase of stroke.

Some limitations of this review should be mentioned. More studies might have used task-specific interventions, but did not report the intervention or the PP5R criteria clear enough to be included in this review. This is partly due to the fact that a clear description of the term "task specific" is lacking. Here we proposed to use a set of criteria, the PP5R criteria, based on existing literature as a first step in defining this term, but a more specific definition is needed which has to be adopted by the field. We are also aware that certain combinations of the PP5R may be questionable whether they reflect task-specific training. In the assessment of task-specific training, we stated that at least four criteria must be fulfilled. It was the first time these criteria were used for task-specificity assessment in a review, which may have led to an underestimation of the task specificity. Studies may have provided a more task-specific training; however, we were not able to verify this from the description of the training. We may have missed relevant results if authors used different words to describe task-specific training in the title, abstract or as keywords. We therefore recommend to use frequently used words such as "task specific" to avoid confusion and elaborate on the provided training to describe why it is task specific in the authors' view. In addition, there was not only a lot of variety in robots used, but also a lot of variety in outcomes, which made it difficult to compare articles. Therefore, we were not able to include all studies in the meta-analyses, which may have introduced bias in our results. We advise to use a core set of outcomes for upper limb stroke research. A core set has been suggested by Kwakkel et al. [73]. They advise to include the FMA-UE and ARAT as upper limb

performance measures. We have seen in this review that the WMFT and MAL are often reported too; however, there was not a standardized way of presenting the results of these tests. We believe that the WMFT is a useful addition to measure the upper limb function since this measure is less focussed on the different grips and more on the arm movements compared with the ARAT. The MAL covers ICF activity and participation, and could therefore be an interesting addition to the measurement set to cover the entire ICF model. However, the MAL seems suitable for patients in the chronic phase of stroke, but might be less suitable in the subacute phase of stroke, since the MAL tasks are often not yet applicable to patients in this phase of stroke. Another difficulty for the meta-analysis was the inconsistent way of presenting the outcomes. Due to a difference in reported FMA-UE outcomes, we were unable to compare all results, despite attempts to contact authors to retrieve missing information. We therefore advise to report outcomes in a standardized way, including pre-test and post-test results and change from baseline including the standard deviations. Another limitation of the study is that we calculated the standard deviation of the change from baseline for some articles since they only provided pre-test and post-test data [34]. Furthermore, since significant differences between pre- and post-test were not always reported, we used a single *t*-test to calculate any significant differences. As a result we may have drawn conclusions in the review and meta-analyses based on incorrect standard deviations or significance. However, the sensitivity analysis showed that the conclusion of the meta-analysis would not change if the correlation coefficient was changed. Furthermore, we excluded CIMT, where this may have been too strict since CIMT is also an often used method for training functional tasks. In retrospect, this did not lead to incorrect exclusion of articles. Articles that performed CIMT and were excluded either did not have a full text available or did not perform task-specific training with the device, but sequenced device training with CIMT. Lastly, the cut off score of the TESTEX quality assessment was not established in the literature, and we therefore determined that 50% of the criteria had to be fulfilled to deem a study of sufficient quality, based on the PEDro cut-off score. An alternative cut-off score might have led to different inclusions or exclusions.

To summarize, the upper limb function in stroke patients can improve after TST-AAD training, although TST-AAD did not appear to be superior over TSUC in the chronic stage. However, in the subacute stage of stroke, TST-AAD seems to be more beneficial than TSUC even after a short training period. We did not find differences between TST-AAD and TSUC for patients in the chronic phase of stroke during the follow-up. TST-AAD can be used to unburden the therapists. Future studies should describe training, device usage and criteria of task specificity (PP5R criteria for assessing task specificity) in a standardized way and use a core set of outcomes to report on upper limb interventions in stroke patients.

Author contributions

SGR searched the databases, CKvds and SGR performed the title selection. KAH and SGR performed the abstract, full text selection and risk of bias assessments. SGR, JMH and CKvds wrote the manuscript. All authors have approved the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the EITHealth under Grant 19094.

ORCID

Samantha G. Rozevink  <http://orcid.org/0000-0002-4992-1578>
 Juha M. Hijmans  <http://orcid.org/0000-0003-3968-9602>
 Corry K. van der Sluis  <http://orcid.org/0000-0003-2141-8214>

References

- [1] Nakayama H, Jorgensen HS, Raaschou HO, et al. Recovery of upper extremity function in stroke patients: the Copenhagen Stroke Study. *Arch Phys Med Rehabil.* 1994; 75(4):394–398.
- [2] Kelly-Hayes M, Robertson JT, Broderick JP, et al. The American Heart Association Stroke outcome classification. *Stroke.* 1998;29(6):1274–1280.
- [3] French B, Thomas L, Leathley M, et al. Does repetitive task training improve functional activity after stroke? A Cochrane systematic review and meta-analysis. *J Rehabil Med.* 2010;42(1):9–15.
- [4] Hubbard IJ, Parsons MW, Neilson C, et al. Task-specific training: evidence for and translation to clinical practice. *Occup Ther Int.* 2009;16(3-4):175–189.
- [5] Bayona NA, Bitensky J, Salter K, et al. The role of task-specific training in rehabilitation therapies. *Top Stroke Rehabil.* 2005;12(3):58–65.
- [6] Platz T, Van Kaick S, Möller L, et al. Impairment-oriented training and adaptive motor cortex reorganisation after stroke: a fTMS study. *J Neurol.* 2005;252(11):1363–1371.
- [7] Hung C-S, Hsieh Y-W, Wu C-Y, et al. The effects of combination of robot-assisted therapy with task-specific or impairment-oriented training on motor function and quality of life in chronic stroke. *Pm R.* 2016;8(8):721–729.
- [8] Schweighofer N, Choi Y, Winstein C, et al. Task-oriented rehabilitation robotics. *Am J Phys Med Rehabil.* 2012;91(11 Suppl 3):S270–S279.
- [9] Van Peppen R, Kwakkel G, Wood-Dauphinee S, et al. The impact of physical therapy on functional outcomes after stroke: what's the evidence? *Clin Rehabil.* 2004;18(8): 833–862.
- [10] Kwakkel G, Kollen BJ, Krebs HI. Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review. *Neurorehabil Neural Repair.* 2008;22(2):111–121.
- [11] Norouzi-Gheidari N, Archambault PS, Fung J. Effects of robot-assisted therapy on stroke rehabilitation in upper limbs: systematic review and meta-analysis of the literature. *J Rehabil Res Dev.* 2012;49(4):479–496.
- [12] Veerbeek JM, Langbroek-Amersfoort AC, Van Wegen EEH, et al. Effects of robot-assisted therapy for the upper limb after stroke: a systematic review and meta-analysis. *Neurorehabil Neural Repair.* 2017;31(2):107–121.
- [13] Bertani R, Melegari C, De Cola MC, et al. Effects of robot-assisted upper limb rehabilitation in stroke patients: a systematic review with meta-analysis. *Neurol Sci.* 2017;38(9): 1561–1569.
- [14] World Health Organization. International classification of functioning, disability and health (ICF). 10th revis. Geneva: World Health Organization; 2011.
- [15] Loureiro RV, Harwin WS, Nagai K, et al. Advances in upper limb stroke rehabilitation: a technology push. *Med Biol Eng Comput.* 2011;49(10):1103–1118.
- [16] Basteris A, Nijenhuis SM, Stienen AHA, et al. Training modalities in robot-mediated upper limb rehabilitation in stroke: a framework for classification based on a systematic review. *J Neuroeng Rehab.* 2014;11:1–15.
- [17] Finley MA, Fasoli SE, Dipietro L, et al. Short-duration robotic therapy in stroke patients with severe upper-limb motor impairment. *J Rehabil Res Dev.* 2005;42(5):683–691.
- [18] Fasoli SE, Krebs HI, Stein J, et al. Effects of robotic therapy on motor impairment and recovery in chronic stroke. *Arch Phys Med Rehabil.* 2003;84(4):477–482.
- [19] Macclellan LR, Bradham DD, Whittall J, et al. Robotic upper-limb neurorehabilitation in chronic stroke patients. *J Rehabil Res Dev.* 2005;42(6):717–722.
- [20] Pila O, Duret C, Laborne FX, et al. Pattern of improvement in upper limb pointing task kinematics after a 3-month training program with robotic assistance in stroke. *J Neuroeng Rehab.* 2017;14:1–10.
- [21] Lo AC, Guarino PD, Richards LG, et al. Robot-assisted therapy for long-term upper-limb impairment after stroke. *N Engl J Med.* 2010;362(19):1772–1783.
- [22] Smart NA, Waldron M, Ismail H, et al. Validation of a new tool for the assessment of study quality and reporting in exercise training studies: TESTEX. *Int J Evid Based Healthc.* 2015;13(1):9–18.
- [23] Maher CG, Sherrington C, Herbert RD, et al. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther.* 2003;83(8):713–721.
- [24] Moseley AM, Herbert RD, Sherrington C, et al. Evidence for physiotherapy practice: a survey of the physiotherapy evidence database (PEDro). *Aust J Physiother.* 2002;48(1): 43–49.
- [25] [Computer program] The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Center, The Cochrane Collaboration;2011.
- [26] van der Lee Johanna H, Beckerman H, Lankhorst GJ, et al. The responsiveness of the action research arm test and the Fugl-Meyer assessment scale in chronic stroke patients. *J Rehabil Med.* 2001;33(3):110–113.
- [27] Fugl-Meyer AR, Jääskö L, Leyman I, et al. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scand J Rehabil Med.* 1975;7(1):13–31.
- [28] Arya KN, Verma R, Garg RK, et al. Meaningful task-specific training (MTST) for stroke rehabilitation: a randomized controlled trial. *Top Stroke Rehabil.* 2012;19(3):193–211.
- [29] Page SJ, Fulk GD, Boyne P. Clinically important differences for the Upper-extremity Fugl-Meyer scale in people with minimal to moderate impairment due to chronic stroke. *Phys Ther.* 2012;92(6):791–798.
- [30] Yozbatiran N, Der-Yeghalian L, Cramer SC. A standardized approach to performing the action research arm test. *Neurorehabil Neural Repair.* 2008;22(1):78–90.
- [31] Lin J-H, Hsu M-J, Sheu C-F, et al. Psychometric comparisons of 4 measures for assessing upper-extremity function in people with stroke. *Phys Ther.* 2009;89(8):840–850.
- [32] Wolf SL, Catlin PA, Ellis M, et al. Assessing wolf motor function test as outcome measure for research in patients after stroke. *Stroke.* 2001;32(7):1635–1639.
- [33] Van Der Lee JH, Beckerman H, Knol DL, et al. Clinimetric properties of the motor activity log for the assessment of arm use in hemiparetic patients. *Stroke.* 2004;35(6): 1410–1414.
- [34] Higgins J, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions version 6.0. John Wiley & Sons; 2019.

- [35] Krebs HI, Mernoff S, Fasoli SE, et al. A comparison of functional and impairment-based robotic training in severe to moderate chronic stroke: a pilot study. *Neurorehabilitation*. 2008;23(1):81–87.
- [36] Bartolo M, Nunzio AD, Sebastiano F, et al. Arm weight support training improves functional motor outcome and movement smoothness after. *Funct Neurol*. 2012;29:15–21.
- [37] Byl NN, Abrams GM, Pitsch E, et al. Chronic stroke survivors achieve comparable outcomes following virtual task specific repetitive training guided by a wearable robotic orthosis (UL-EXO7) and actual task specific repetitive training guided by a physical therapist. *J Hand Ther*. 2013;26(4):343–352.
- [38] Casadio M, Giannoni P, Morasso P, et al. A proof of concept study for the integration of robot therapy with physiotherapy in the treatment of stroke patients. *Clin Rehabil*. 2009;23(3):217–228.
- [39] Colombo R, Pisano F, Micera S, et al. Assessing mechanisms of recovery during robot-aided neurorehabilitation of the upper limb. *Neurorehabil Neural Repair*. 2008;22(1):50–63.
- [40] Connelly L, Jia Y, Toro ML, et al. A pneumatic glove and immersive virtual reality environment for hand rehabilitative training after stroke. *IEEE Trans Neural Syst Rehabil Eng*. 2010;18(5):551–559.
- [41] Daunoraviciene K, Adomaviciene A, Grigonyte A, et al. Effects of robot-assisted training on upper limb functional recovery during the rehabilitation of poststroke patients. *Technol Health Care*. 2018;26:533–542.
- [42] Tomić TJD, Savić AM, Vidaković AS, et al. ArmAssist robotic system versus matched conventional therapy for poststroke upper limb rehabilitation: a randomized clinical trial. *Biomed Res Int*. 2017;2017:7659893.
- [43] Fischer HC, Triandafilou KM, Thielbar KO, et al. Use of a portable assistive glove to facilitate rehabilitation in stroke survivors with severe hand impairment. *IEEE Trans Neural Syst Rehabil Eng*. 2016;24(3):344–351.
- [44] Fluet GG, Merians AS, Qiu Q, et al. Does training with traditionally presented and virtually simulated tasks elicit differing changes in object interaction kinematics in persons with upper extremity hemiparesis? *Top Stroke Rehabil*. 2015;22(3):176–184.
- [45] Housman SJ, Scott KM, Reinkensmeyer DJ. A randomized controlled trial of gravity-supported, computer-enhanced arm exercise for individuals with severe hemiparesis. *Neurorehabil Neural Repair*. 2009;23(5):505–514.
- [46] Huang T-Y, Pan L-L, Yang W-W, et al. Biomechanical evaluation of Three-Dimensional printed dynamic hand device for patients with chronic stroke. *IEEE Trans Neural Syst Rehabil Eng*. 2019;27(6):1246–1252.
- [47] Klamroth-Marganska V, Blanco J, Campen K, et al. Three-dimensional, task-specific robot therapy of the arm after stroke: a multicentre, parallel-group randomised trial. *Lancet Neurol*. 2014;13(2):159–166.
- [48] Lamercy O, Dovat L, Yun H, et al. Effects of a robot-assisted training of grasp and pronation/supination in chronic stroke: a pilot study. *J Neuroeng Rehabil*. 2011;8:63.
- [49] Page S, Hill V, White S. Portable upper extremity robotics is as efficacious as upper extremity rehabilitative therapy: a randomized controlled pilot trial. *Arch Phys Med Rehabil*. 2012;93(10):e21.
- [50] Page S, Griffin C, White S. Efficacy of myoelectric bracing in moderately impaired stroke survivors: a randomized, controlled trial. *J Rehabil Med*. 2020;52:rm00017.
- [51] Takahashi CD, Der-Yeghiaian L, Le V, et al. Robot-based hand motor therapy after stroke. *Brain*. 2008;131(2):425–437.
- [52] Thielbar KO, Triandafilou KM, Fischer HC, et al. Benefits of using a voice and EMG-driven actuated glove to support occupational therapy for stroke survivors. *IEEE Trans Neural Syst Rehabil Eng*. 2017;25(3):297–306.
- [53] Timmermans AAA, Lemmens RJM, Monfrance M, et al. Effects of task-oriented robot training on arm function, activity, and quality of life in chronic stroke patients: a randomized controlled trial. *J Neuroeng Rehabil*. 2014;11:45–11.
- [54] Kwakkel G, Van Peppen R, Wagenaar RC, et al. Effects of augmented exercise therapy time after stroke: a meta-analysis. *Stroke*. 2004;35(11):2529–2539.
- [55] Oujamaa L, Relave I, Froger J, et al. Rehabilitation of arm function after stroke. Literature review. *Ann Phys Rehabil Med*. 2009;52(3):269–293.
- [56] Timmermans AAA, Seelen HAM, Willmann RD, et al. Technology-assisted training of arm-hand skills in stroke: concepts on reacquisition of motor control and therapist guidelines for rehabilitation technology design. *J Neuroeng Rehabil*. 2009;6:1.
- [57] Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet*. 2011;377(9778):1693–1702.
- [58] Starkey ML, Schwab ME. How plastic is the brain after a stroke? *Neuroscientist*. 2014;20(4):359–371.
- [59] Cramer SC. Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Ann Neurol*. 2008;63(3):272–287.
- [60] Kreisel SH, Hennerici MG, Bänzner H. Pathophysiology of stroke rehabilitation: the natural course of clinical recovery, use-dependent plasticity and rehabilitative outcome. *Cerebrovasc Dis*. 2007;23(4):243–255.
- [61] Teasell R, Mehta S, Pereira S, et al. Time to rethink long-term rehabilitation management of stroke patients. *Top Stroke Rehabil*. 2012;19(6):457–462.
- [62] Woytowicz EJ, Rietschel JC, Goodman RN, et al. Determining levels of upper extremity movement impairment by applying a cluster analysis to the Fugl-Meyer assessment of the upper extremity in chronic stroke. *Arch Phys Med Rehabil*. 2017;98(3):456–462.
- [63] Stein J, Krebs HI, Frontera WR, et al. Comparison of two techniques of robot-aided upper limb exercise training after stroke. *Am J Phys Med Rehabil*. 2004;83(9):720–728.
- [64] Nijenhuis SM, Prange GB, Amirabdollahian F, et al. Feasibility study into self-administered training at home using an arm and hand device with motivational gaming environment in chronic stroke. *J Neuroeng Rehabil*. 2015;12:1–12.
- [65] Buurke JH, Nene AV, Kwakkel G, et al. Recovery of gait after stroke: what changes? *Neurorehabil Neural Repair*. 2008;22(6):676–683.
- [66] Farina D, Zennaro D, Pozzo M, et al. Single motor unit and spectral surface EMG analysis during low-force, sustained contractions of the upper trapezius muscle. *Eur J Appl Physiol*. 2006;96(2):157–164.
- [67] Buma F, Kwakkel G, Ramsey N. Understanding upper limb recovery after stroke. *Restor Neurol Neurosci*. 2013;31(6):707–722.
- [68] Kwakkel G, Kollen B, Lindeman E. Understanding the pattern of functional recovery after stroke: facts and theories. *Restor Neurol Neurosci*. 2004;22(3-5):281–299.

- [69] Szafarski JP, Page SJ, Kissela BM, et al. Cortical reorganization following modified constraint-induced movement therapy: a study of 4 patients with chronic stroke. *Arch Phys Med Rehabil.* 2006;87(8):1052–1058.
- [70] Liepert J, Bauder H, Miltner WHR, et al. Treatment-induced cortical reorganization after stroke in humans. *Stroke.* 2000; 31(6):1210–1216.
- [71] Hlušík P, Mayer M. Paretic hand in stroke: from motor cortical plasticity research to rehabilitation. *Cogn Behav Neurol.* 2006;19:34–40.
- [72] Thirumala P, Hier DB, Patel P, et al. Motor recovery after stroke: lessons from functional brain imaging. *Neurol Res.* 2002;24(5):453–458.
- [73] Kwakkel G, Lannin NA, Borschmann K, et al. Standardized measurement of sensorimotor recovery in stroke trials: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *Int J Stroke.* 2017; 12(5):451–461.