

REVIEW ARTICLE

## SYSTEMATIC REVIEW OF OUTCOME MEASURES OF WALKING TRAINING USING ELECTROMECHANICAL AND ROBOTIC DEVICES IN PATIENTS WITH STROKE

Christian Geroin, PT<sup>1\*</sup>, Stefano Mazzoleni, PhD<sup>3\*</sup>, Nicola Smania, MD<sup>1,2</sup>, Marialuisa Gandolfi, MD, PhD<sup>1</sup>, Donatella Bonaiuti, MD<sup>4</sup>, Giulio Gasperini, MD<sup>5</sup>, Daniele Munari, PT<sup>1</sup>, Patrizio Sale, MD, PhD<sup>6</sup>, Andreas Waldner, MD<sup>7</sup>, Raffaele Spidalieri, MD<sup>8</sup>, Federica Bovolenta, MD<sup>9</sup>, Alessandro Picelli, MD<sup>1</sup>, Federico Posteraro, MD<sup>10</sup>, Franco Molteni, MD<sup>5</sup>, Marco Franceschini, MD<sup>6</sup> and the Italian Robotic Neurorehabilitation Research Group (IRNRG)

*From the <sup>1</sup>Neuromotor and Cognitive Rehabilitation Research Centre (CRRNC), Department of Neurological and Movement Sciences, University of Verona, <sup>2</sup>Neurological Rehabilitation Unit Azienda Ospedaliera-Universitaria Integrata Verona, <sup>3</sup>The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, <sup>4</sup>Physical Medicine and Rehabilitation Department, S. Gerardo Hospital, Monza, <sup>5</sup>Department of Rehabilitation Medicine, Ospedale Valduce, Villa Beretta, Costamasnaga, Lecco, <sup>6</sup>Department of Rehabilitation IRCCS San Raffaele Pisana, Rome, <sup>7</sup>Department of Neurological Rehabilitation, Private Hospital Villa Melitta, Bolzano, <sup>8</sup>Istituto di Riabilitazione Neurologica "Madre Della Divina Provvidenza" di Agazzi, Arezzo, <sup>9</sup>Medicine Rehabilitation NOCSAE Hospital AUSL of Modena, Modena and <sup>10</sup>Neurological Rehabilitation Unit, Auxilium Vitae Rehabilitation Center, Volterra, Italy. \*Both authors contributed equally to this work.*

**Objective:** The aim of this systematic review was to identify appropriate selection criteria of clinical scales for future trials, starting from those most commonly reported in the literature, according to their psychometric properties and International Classification of Functioning, Disability and Health (ICF) domains.

**Data sources:** A computerized literature research of articles was conducted in MEDLINE, EMBASE, CINALH, PubMed, PsychINFO and Scopus databases.

**Study selection:** Clinical trials evaluating the effects of electromechanical and robot-assisted gait training trials in stroke survivors.

**Data extraction:** Fifteen independent authors performed an extensive literature review.

**Data synthesis:** A total of 45 scales was identified from 27 studies involving 966 subjects. The most commonly used outcome measures were: Functional Ambulation Category (18 studies), 10-Meter Walking Test (13 studies), Motricity Index (12 studies), 6-Minute Walking Test (11 studies), Rivermead Mobility Index (8 studies) and Berg Balance Scale (8 studies). According to the ICF domains 1 outcome measure was categorized into Body Function and Structure, 5 into Activity and none into Participation.

**Conclusion:** The most commonly used scales evaluated the basic components of walking. Future studies should also include instrumental evaluation. Criteria for scale selection should be based on the ICF framework, psychometric properties and patient characteristics.

**Key words:** stroke; lower limb; rehabilitation; motor recovery; robot; training; therapy; physiotherapy; function; study; robot-assisted, trial.

*Correspondence address: Christian Geroin, Neuromotor and Cognitive Rehabilitation Research Centre (CRRNC), Department of Neurological, and Movement Sciences, University of Verona, 37134 Verona, Italy. E-mail: christian.geroin@univr.it*

Accepted Jun 17, 2013; Epub ahead of print Oct 17, 2013

### INTRODUCTION

Stroke is a leading cause of disability (1, 2). Among areas with population-based studies, the overall age-standardised incidence of stroke in people aged  $\geq 55$  years range from 4.2 to 11.7 per 1,000 person-years (1). Approximately 64% of stroke survivors have persisting sensorimotor deficits leading to progressive upper and lower limb disability (3), which restricts their autonomy in activities of daily living (ADLs). Recovery of walking is one of the main objectives in stroke rehabilitation, which contributes to an improvement in independence (4).

Conventional rehabilitation has been proven, to some extent, to be effective in improving walking function; however, it often requires great physical effort by physiotherapists (4). In recent years, several innovative technologies and strategies have been proposed to overcome this difficulty and improve walking function (4–6). According to the modern concept of task-specific training, electromechanical and robotic-assisted gait training, in combination with conventional rehabilitation, has been shown to be feasible and effective to improve walking in stroke survivors (4, 5), even facilitating repetitive practice of gait-like movement in individuals who are wheelchair users. Although research regarding these neurorehabilitation approaches is growing, literature concerning specific outcome measures is scant (7).

The evaluation of treatment outcomes is a key factor in both clinical rehabilitation practice and research settings, but there is no agreement on the most appropriate modality to select outcome measures (7–10). Three main limitations can be identified. First, a large number of instruments is available, but they have poor psychometric properties. Secondly, there is no shared consensus on specific clinical outcome measures that should be used to assess the effects of electromechanical and robot-assisted gait training trials (ERAGTT). Finally, the outcome measures regarding the evaluation of recovery of function and compensation adaptation processes, which strongly affect the patient's involvement in ADLs, are often unclear and misinterpreted (11, 12).

Choosing a suitable scale to assess sensorimotor recovery is a challenging issue in rehabilitation, given that several constraints could interfere with their appropriate selection (10). For instance, the domain to be measured (e.g. function, activity, quality of life), clinical area (e.g. neurological, geriatric), setting (e.g. hospital, community, home), as well as psychometric properties (e.g. reliability, validity, responsiveness) could interfere with the selection of the most appropriate outcome measures.

The aim of this systematic review is to identify appropriate selection criteria of clinical scales for future trials, starting

from those most commonly used in the literature, according to their psychometric properties and International Classification of Functioning, Disability and Health (ICF) domains.

## MATERIAL AND METHODS

The systematic review was performed by the authors in 3 stages, as described below, according to the methodology reported by Sivan et al. (7).

*Stage 1: Search for clinical trials involving electromechanical and robot-assisted gait training in patients after stroke and determine the outcome measures used in each trial*

*Data sources.* A search of MEDLINE, EMBASE, CINALH, PubMed, PsychINFO and Scopus databases was performed to identify relevant ERAGTT. The keywords used were: stroke, lower limb, rehabilitation, motor recovery, robot, training, therapy, physiotherapy, function, study, robot-assisted and trial. From the initial search, all abstracts were reviewed.

*Study selection.* Inclusion criteria were: (i) studies published from January 2000 to January 2012; studies involving participants with diagnosis of a stroke; (ii) lower limb exercise assisted by a robot device. A robotic device was defined as any technology able to assist the patient's limb movement for therapeutic exercises, to support the therapist during administration of programmable and customized rehabilitation programmes and composed of mechanical structure with actuators and energy supply; (iii) at least one scale used in the study.

Table I. Definition and standard values for the evaluation criteria. (Modified with permission from ref 7)

Properties	Definition of the properties	Standard values
Reliability	Reproducibility of an outcome measure is defined as the amount of the score that includes information about the characteristic of interest opposite to measurement error (10). Reliability can be evaluated in 3 basic ways: (i) test-retest reliability; (ii) inter-rater reliability; and (iii) internal consistency reliability (10).	Test-retest or inter-rater reliability (Icc; kappa statistics): excellent: $\geq 0.75$ ; adequate: 0.4–0.74; poor: $\leq 0.40$ . A minimum test-retest reliability of 0.90 is recommended whether the measure is performed during the ongoing progress of a subject undergoing treatment (15). Internal consistency (split-half or Cronbach's $\alpha$ statistics): excellent: $\geq 0.80$ ; adequate: 0.70–0.79; poor: $< 0.70$ (15).
Validity	Validity is the faculty of a scale to measure what it is intended to measure. Many types of validity exist in literature, e.g. face, content, discriminative, convergent, predictive, and criterion. The most important are criterion and predictive validity (10).	Construct/convergent and concurrent correlations: excellent: $\geq 0.60$ ; adequate: 0.31–0.59; poor: $\leq 0.30$ . ROC analysis – AUC: excellent: $\geq 0.90$ ; adequate: 0.70–0.89; poor: $< 0.70$ . No agreement on ideal values by which to judge sensitivity and specificity as a validity index (15).
Responsiveness	Responsiveness is sensitivity to changes within patients over time, which could be indicative of therapeutic effects. Minimal clinically important difference (MCID) is the smallest score difference in the domain of interest that patients perceive as beneficial (10). Floor and ceiling indicate limits to the range of evident modification beyond which no further improvement or worsening can be detected (10).	Sensitivity to change: excellent: with standardized effect sizes: $< 0.5$ = small; 0.5–0.8 moderate; $\geq 0.8$ = large. Further available methods are: Standardized Response Mean (SRM), ROC Analysis – Area Under Curve (AUC), Statistical Significance $p$ -value, correlation values of observed change compared to change in other scales, MCID described as a score value (7). Adequate: evidence of moderate/less change than expected; contradictory evidence. Poor: feeble evidence based solely on $p$ -values (statistical significance). Floor/ceiling effects: excellent: no floor or ceiling effects; adequate: floor and ceiling effects $< 20\%$ ; poor: $> 20\%$ (15).
Acceptability	Acceptability can be divided into respondent and administrative burden. Respondent refers to whether the length and content are acceptable to the intended participants (e.g. stroke individuals). Administrative refers to whether the tool is user-friendly, easy to understand and cheap (7).	Respondent burden: optimal – time to administration less than 15 min and easy to understand; adequate – longer or some problems of acceptability; poor – problems of acceptability and lengthy (7). Administrative burden: optimal when score is immediately obtained and easy to understand; adequate when score requires interpretation by computer; and poor when score is complex and expensive to be detected (7).

ROC: receiver operator characteristic; AUC: area under curve; ICC: intraclass correlation coefficient.

The exclusion criteria were: (i) studies involving a robotic orthosis device; (ii) studies enrolling only healthy volunteers; and (iii) articles published in languages different from English.

**Data extraction.** Multiple independent investigators performed the article selection as follows: 15 investigators carried out an extensive literature review and selected the studies according to the inclusion criteria; NS, FP, GG, FM and PS independently read in detail all the selected articles; MG, AW, RS, FB and AP reviewed the same articles and listed the scales used; DB, DM, MF, CG, SM performed a review based on the psychometric properties of the different scales; CG, SM and MG drafted the manuscript. Disagreements were resolved by discussion between authors. All authors have read, edited and agreed on the contents of the manuscript.

In this review the term “scale” was used to define the assessment instrument used as a discriminative or predictive tool in ERAGTT. Discriminative scales are used to cluster patients into homogeneous groups for treatment studies (10). Predictive scales are used to predict how the motor recovery will evolve over time. The term “outcome measure” was used to define the evaluative instruments that reflect clinically important changes after intervention (10). Evaluative instruments are used to estimate the quantity of longitudinal change in an individual or group of patients who underwent the rehabilitation intervention (10).

### Stage 2. List and classify the scales collected during stage 1 according to the ICF domains

The content of each scale identified in Step 1 was classified in terms of the ICF categories, according to literature classification and specific website research (9, 13, 14). When necessary, the scale classification was discussed between authors. Three categories were identified as follows:

- **Body functions and structures:** functions refer to physiological functions of body systems including psychological. Structures are anatomical parts or regions of the body and their components. Impairments are defined as problems or disorders in body function or structure (9).
- **Activity:** activity refers to execution of a task by an individual. Limitations of a task are defined as difficulties an individual might experience in completing a given activity (15).
- **Participation:** involvement of an individual in a life situation. Restrictions to participation describe difficulties experienced by the individual in a life situation or role (16).
- **Contextual factors:** which include environmental and personal factors that may influence the relationship among different factors (16).

### Stage 3. Describe the measurement properties of the identified scales in patients after stroke

A literature search of the psychometric properties of each scale was performed. The reliability, validity and responsiveness of each scale were investigated. The score for each property was identified as high or excellent (+++), moderate (++) or poor (+) (16, 17).

Moreover, minimal clinically important difference (MCID), floor and ceiling effect, time of administration and level of measurement (nominal, ordinal, interval and ratio) for each scale were evaluated. Table I describes the definition and standards values of the psychometric properties considered. A further classification of the scales used in the trials according to phase of disease was performed.

## RESULTS

### Stage 1

A total of 27 studies published from 2000 to 2012 (involving 966 subjects) fulfilled the inclusion criteria for the review. A total of 45 scales was identified. The list of the scales used in these studies and the corresponding abbreviations are provided in Table II. Details regarding the type of electromechanical

or robot device, authors, number and type of patients and the scales used are provided in Table III. The most common outcome measures used were: Functional Ambulation Category (FAC; 18 studies); 10-Meter Walking Test (10MWT; 13 studies); Motricity Index (MI; 12 studies); 6-Minute Walking Test (6MinWT; 11 studies); Rivermead Mobility Index (RMI; 8 studies); and Berg Balance Scale (BBS; 8 studies). The scales reported in Table III considered as “others” represent a mix of discriminative, evaluative and predictive scales.

Table II. Abbreviations for the scales

Abbreviation	Scales
2MinWT	2-Minute Walking Test (18)
3MinWT	3-Minute Walking Test (19)
5MWT	5-Meter Walking Test (20)
6MinWT	6-Minute Walking Test (21)
8MWT	8-Meter Walking Test (19)
10MWT	10-Meter Walking Test (22)
AS	Ashworth Scale (23)
BBS	Berg Balance Scale (24)
BI	Barthel Index (25)
BMI	Body mass index (26)
CES-D	Center for Epidemiological Studies-Depression Scale (27)
CNS	Canadian Neurological Scale (28)
EMS	Elderly Mobility Scale (29)
ESS	European Stroke Scale (30)
FAC	Functional Ambulation Category (8)
FAI	Frenchay Activities Index (31)
FIM	Functional Independence Measure (32)
FM motor	Fugl-Meyer Motor Subscale (33)
FMA	Fugl-Meyer Assessment of Sensorimotor Recovery After Stroke (33)
HR	Heart rate
LLFDI	Late Life Function and Disability Instrument (34)
MAS	Modified Ashworth Scale (23)
MEFAP	Modified Emory Functional Ambulation Profile (20)
MI	Motricity Index (35)
MMAS	Modified Motor Assessment Scale (36)
MMSE	Mini Mental State Examination (37)
MoAS	Motor Assessment Scale (38)
MRC	Medical Research Council (39)
MRS	Modified Ranking Scale (40)
NIHSS	National Institutes of Health Stroke Scale (41)
PROM	Passive Range of Movement
RMAS	Rivermead Motor Assessment Scale (42)
RMI	Rivermead Mobility Index (43)
RPE	Borg Scale of Perceived Exertion (44)
RS	Rankin Scale (45)
SAS	Stroke Activities Scale (46)
SF-36	Short Form Health Survey (47)
SPPB	Short Physical Performance Battery (48)
SSS	Scandinavian Stroke Scale (49)
ST	Step Test (50)
TBS	Tinetti Balance Scale (51)
TCT	Trunk Control Test (35)
TGS	Tinetti Gait Scale (20)
TMS	Toulouse Motor Scale (52)
TUG	Timed Up and Go Test (20)
<i>Instrumental measures</i>	
JK	Joint Kinematic
STGP	Spatio-temporal Gait Parameters

Table III. Scales used in ERAGTT (classified by number of studies and year of publication)

Electromechanical/ robotic device	Reference	n	Type of patients	Most commonly used outcome measures						
				FAC	10MWT	MI	6MinWT	RMI	BBS	Others
G-EO GT1	Hesse et al., 2010 (53)	1	Subacute	*		*		*		BI
	Conesa et al, 2012 (54)	103	Subacute	*	*					TBS, TGS
	Morone et al., 2011 (55)	48	Subacute	*	*	*	*	*		AS, BI, CNS, MMSE, RS, TCT
	Geroin et al., 2011 (56)	30	Chronic	*	*	*	*	*		ESS, MMSE, STGP, MAS
	Peurala et al., 2009 (57)	56	Subacute	*	*		*	*		BI, BMI, HR, MMAS, MRS, RMAS, RPE, SSS
	Maple et al., 2008 (58)	54	Subacute	*		*			*	5MWT, BI, EMS, FIM, MMSE
	Pohl et al., 2007 (59)	155	Subacute	*	*	*	*	*		BI, MRC, PROM
	Dias et al., 2007 (52)	40	Chronic	*	*	*	*	*	*	BI, FM motor, MMSE, ST, TMS, TUG, MAS
	Tong et al., 2006 (60)	46	Subacute	*		*			*	MMSE, 5MWT, EMS, FIM, BI
	Peurala et al., 2005 (61)	45	Chronic		*		*			FIM, MRC, MMAS, RPE, SSS, postural sway (Force Plate), HR, FAC, MAS
LK	Werner et al., 2002 (62)	30	Subacute	*	*					RMAS, BI, MAS
	Hesse et al., 2001 (63)	14	Chronic	*	*					RMAS, EMG, STGP, MAS
	Hesse et al., 2000 (64)	2	Subacute	*						RMAS, MAS
	Hesse et al., 2000 (65)	2	Subacute	*						RMAS, MAS
	Chang et al., 2011 (66)	37	Subacute	*		*				FM motor, AC, CR, VR
	Magagnin et al., 2010 (67)	5	Chronic			*				BI, FIM, TCT, ECG
	Lewek et al., 2009 (68)	19	Chronic							MMSE, STGP, JK
	Hidler et al., 2009 (69)	63	Subacute	*			*	*	*	5MWT, FAI, MoAS, NIHSS, SF-36, MMSE, CES-D, STGP
	Schwartz et al., 2009 (70)	67	Subacute	*	*					2MWT, FIM, NIHSS, SAS, TUG
	Westlake et al., 2009 (71)	16	Chronic				*		*	FM motor, LLFDI, SPPB, STGP
AA LH CaLT	Hornby et al., 2008 (72)	48	Chronic				*		*	CES-D, FAI, MEFAP, MRC, MMSE, SF-36, STGP, MAS
	Mayr et al., 2007 (73)	16	Mixed		*	*	*			AS, MRC, RMAS
	Krewer et al., 2007 (74)	10	Mixed							BMI, HR, Energy expenditure
	Husemann et al., 2007 (75)	30	Acute	*	*	*				BI, FAC, MRC, STGP, MAS
	Fisher et al., 2011 (19)	20	Mixed							3MWT, 8MWT, MMSE, TBS
	Freivogel et al., 2009 (76)	2	Chronic	*	*	*		*	*	PROM, MAS
	Wu et al., 2011 (77)	7	Chronic				*		*	STGP, MMSE

\*Used in trial; G-EO: G-EO System; GT1: Gait-Trainer GT1; LK: Lokomat; AA: Autoambulator; LH: LokoHELP; FAC: Functional Ambulation Category; 10MWT: 10-Meter Walk Test; MI: Motricity Index; 6MinWT: 6-Minute Walk Test; RMI: Rivermead Mobility Index; BBS: Berg Balance Scale; STGP: Spatiotemporal Gait Parameters; JK: Joint Kinematic; AC: aerobic capacity; CR: cardiovascular response; VR: ventilatory response; EMG: electromyography; ECG: electrocardiography; CaLT: novel cable-driven robotic gait training system. For other abbreviations, see Table II.

Stage 2

Each scale was classified into a single ICF domain, as shown in Fig. 1. Eighteen scales were classified into the body function domain, 24 scales into the activity and 3 into the participation.

Stage 3

The psychometric properties of the most commonly used outcome measures, based on the purpose of the measurement (10), are described in Table IV, whereas the levels of measurement according to Stevens (80) are reported in Table V (29 scales were ordinal, 12 ratio and 4 nominal). The classification of the scales used in the trials according to phase of disease is reported in Table VI (10 in the acute, 6 in chronic and 29 scales in both phases).

DISCUSSION

Our results show that FAC, 10MWT, MI, 6MinWT, RMI and BBS were the most used commonly outcome measures in ERAGTT. As regards ICF classification, they mainly belong to the activity domain (FAC, 10MWT, 6MinWT, RMI, and BBS) and only one to body function and structure (MI). No scale belonged to participation category (Fig. 1).

ICF classification

Body Function and Structures. The function level is an essential part of the assessment process. However, this level alone cannot provide information on whether the improvements are related to recovery of function or to compensation (12). The scales included in this classification often provide specific information

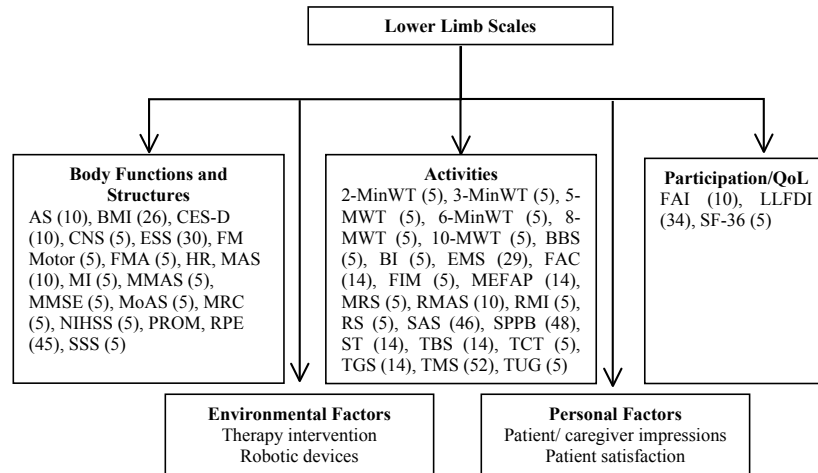


Fig. 1. International Classification of Functioning, Disability and Health (ICF) categorization of scales used in studies on the effects of rehabilitation treatments using electromechanical and robotic devices. “()”: ICF classification reference for each scale. For abbreviations of the scales, see Table II.

regarding the quantity of movement performed by a subject, but not about the quality of movement needed to distinguish between the 2 different recovery processes (12). From a clinical point of view this represents a substantial weakness. Previous studies have not distinguished the 2 different processes in the selection of scales. Thus, the impairment scales should be accompanied with data about the quality of movement, as, for instance, provided by a dynamic electromyography (EMG) evaluation and gait analysis.

The main finding of this review is that the MI is the most widely used and reliable scale to evaluate body function and structure. Thus, post-stroke strength training represents an important part of a rehabilitation programme. The validity of the MI in the evaluation of lower limb muscle strength is also confirmed by instrumental strength evaluations, such as the dynamometer (79). However, the MI does not provide information regarding quality of motor performance and other associated phenomena (35), which could be important to evaluate specific

ERAGTT effects. It is also noteworthy that the MI includes one sub-item (ankle dorsiflexion) that has been considered as a potential predictive factor of lower limb motor recovery (81).

*Activity.* The activity level is an essential part of the assessment process as well. However, the activity level alone cannot provide information on whether the improvements are related to the recovery of function or to compensation (12) because the term “limitation in activity” refers to one’s difficulty in completing a given task. Thus, the activity scales should be accompanied by quality of movement assessment, as previously discussed (12).

Furthermore, the activity recovery may not be necessarily correlated with improvements in ADLs, because an individual may improve in the activity domain with a scarce impact in their level of ADL independence in their social environment. In this context, the assessment of activity should be associated with participation scales, a crucial issue that should be considered in future studies (7).

Table IV. Psychometric properties of the most commonly used outcome measures in electromechanical and robot-assisted gait training trials

Characteristics	FAC	10MWT	MI	6MinWT	RMI	BBS
Time taken (min)	1	5	20	6	4	10–15
Number of items	1	1	6	n/a	15	14
Type	1p	Timed	0–33p	Meter	2p	4p
Score range	1–6	Varies	0–33	Varies	0–15	0–56
Test-retest reliability	+++	+++	n/a	+++	+++	+++
Inter-rater reliability	+++	+++	+++	+++	+++	+++
Construct validity	+++	+++	+++	+++	+++	+++
Responsiveness	++	+++	n/a	n/a	+++	+++
MCID	n/a	0.16 m/s	n/a	50 m	3	n/a
Floor effect	n/a	n/a	n/a	n/a	adeq	adeq
Ceiling effect	n/a	poor	n/a	n/a	adeq	adeq
Burden	adeq	adeq	adeq	adeq	adeq	adeq
References	8	20, 78	35, 79	20	15, 20, 43	24

Scoring criteria as define in Table I. For abbreviations, see Table II.

+++High/excellent; ++moderate; +low/poor; n/a: no available evidence yet; adeq: adequate (acceptable) floor/ceiling effect/burden; poor: poor (unacceptable) floor/ceiling effect/burden; nil: minimal/no burden; MCID: minimal clinically important difference.

Table V. Scales classified according to levels of measurement (80)

Scales	Nominal	Ordinal	Interval	Ratio
10MWT				*
2MinWT				*
3MinWT				*
5MWT				*
6MinWT				*
8MWT				*
AS		*		
BBS		*		
BI		*		
BMI				*
CES-D		*		
CNS		*		
EMS		*		
ESS		*		
FAC		*		
FAI		*		
FIM		*		
FM motor		*		
FMA		*		
HR				*
LLFDI		*		
MAS		*		
MEFAP				*
MI		*		
MMAS		*		
MMSE		*		
MoAS		*		
MRC		*		
MRS		*		
NIHSS		*		
PROM				*
RMAS		*		
RMI		*		
RPE		*		
RS		*		
SAS		*		
SF-36	*			
SPPB	*			
SSS	*			
ST				*
TBS		*		
TCT		*		
TGS		*		
TMS	*			
TUG				*
<i>Instrumental measures</i>				
JK				*
STGP				*

For abbreviations, see Table II.

The most commonly used scales to evaluate at the activity level in ERAGTT were the FAC, 10MWT, 6MinWT, RMI, and BBS.

The 10MWT is widely used to evaluate speed of walking (22). Velocity, in fact, is a component of walking that allows an individual to move within the home environment and the community (e.g. cross a street). Many individuals post-stroke are sedentary, which, when combined with normal ageing, predisposes them to increased functional deficits and declined activity tolerance (82).

Table VI. Scales classified according to phase of disease used in the studies

Scales	Phase of disease			
	Acute		Chronic	
	Severe impairment	Moderate impairment	Severe impairment	Moderate impairment
2MinWT	*	*		
3MinWT		*		*
5MWT	*	*		
6MinWT	*	*	*	*
8MWT		*		*
10MWT	*	*	*	*
AS	*	*	*	*
BBS	*	*	*	*
BI	*	*	*	*
BMI	*	*	*	*
CES-D		*		*
CNS	*			
EMS	*			
ESS				*
FAC	*	*	*	*
FAI		*		*
FIM	*	*	*	*
FM motor	*			*
FMA				
HR	*	*	*	*
LLFDI				*
MAS	*		*	*
MEFAP				*
MI	*	*	*	*
MMAS	*	*	*	*
MMSE	*	*		*
MoAS		*		
MRC	*	*	*	*
MRS	*	*		
NIHSS	*	*		
PROM	*		*	
RMAS	*	*	*	*
RMI	*	*	*	*
RPE	*	*	*	*
RS	*			
SAS	*	*		
SF-36		*		*
SPPB				*
SSS	*	*	*	*
ST				*
TBS	*	*		*
TCT	*		*	
TGS	*	*		
TMS		*		*
TUG	*	*		*

Severe impairment – Functional Ambulation Category  $\leq 2$ : the patient is not able to ambulate independently. Moderate impairment – Functional Ambulation Category  $\geq 3$ : the patient is able to ambulate with verbal supervision, without physical contact. For abbreviations, see Table II.

Many studies used the 6MinWT to evaluate endurance of walking. Patients after stroke have shown a reduction in both strength and cardiorespiratory fitness (83). Thus, improving endurance of gait is one of the most important aims that should always be considered during ERAGTT. Correlations between improved walking endurance and decreased disability post-stroke are reported in several trials (4).

With regards to the assessment of mobility, the most widely used scale was the RMI. Mobility is one of the most important objectives in rehabilitation because its impairment has deleterious effects on ADLs and quality of life. A recent study showed that RMI can be used to predict the length of institutional stay for people with stroke within 5 days after stroke onset (84).

Finally, the BBS was the scale mainly used to evaluate balance, which is a very important skill in order to prevent falls and improve gait performance.

*Participation.* The participation domain, which represents one of the most challenging research issues in neurorehabilitation, has been partially neglected when selecting ERAGTT scales. Up to now, few studies have analysed the impact of the ERAGTT on improving individuals' involvement in real-life situations, defined as participation (12).

During the rehabilitation period, robotic devices can be used to improve body functions/activities and to provide a quantitative assessment. Furthermore, the therapist should use these functional improvements to promote generalization processes in order to increase independence in ADLs. Future studies should consider this aspect as the ultimate goal of stroke rehabilitation, to discharge patients as functional community-dwelling adults.

It is noteworthy that this review process highlighted other important issues that require further discussion. In particular, (i) the time needed to perform the assessment, (ii) the psychometric properties of the scales, (iii) the phase of disease in which they were administered; and, finally, (iv) a proposal of battery of tests for future studies.

#### *Time of scale administration*

Our findings showed that the most commonly used scales are simple and do not require more than 20 min to administer (Table IV). The time required to administer a scale is an important feature. Indeed, many scales often require a long administration time, rendering them inappropriate in some contexts, such as in busy outpatient clinics (85).

#### *Psychometric properties*

The psychometric properties, such as reliability, validity, responsiveness, sensibility and MCID (10), represent important factors when selecting the most appropriate outcome measures (Table IV). They are no fixed scale properties, but they depend on the type of disease, on the phase of illness and on the population studied (10).

Almost all scales use in ERAGTT, except for the MI, are reliable. Reliability is a very important property for patient-based outcome measures in clinical trials. It is essential to establish that any change observed could be due to the intervention itself and is not related to any other problem in the measuring process. However, it does not yield any information about scale validity (10), such as, for instance, construction validity that refers to whether a scale measures or correlates with another measure to provide a basis for comparison (16). It is important to note that all of the most commonly used outcome measures have good construct validity.

As for responsiveness, the FAC, BBS, RMI, and 10MWT presented a large responsiveness value, while it was not reported for 6MinWT and MI. Responsiveness is the sensitivity of a scale to change within patients over time. One of the main limitations on the responsiveness of an instrument regards to the ceiling and floor effects. These data are not reported for FAC, 6MinWT and MI (Table IV).

MCID, which is defined as the smallest difference score in the domain of interest, which the individual feels as beneficial, was found only for the 6MinWT and 10MWT.

To conclude, the results showed that several properties of the scales are not currently available. Further studies are needed to obtain the missing properties during different phases of disease.

#### *The most-used outcome measures according to the severity and phase of disease*

A further analysis of these studies was performed to evaluate the severity and phase of disease where the scales were used. Our intention is to provide an overall perspective of the scales used in ERAGTT, classified according to the ambulation independence by the FAC scale as a benchmark, due to its large diffusion. We believe that such classification could help clinicians to choose the most appropriate scale during clinical practice.

Based on the scale's clinical significance, we considered patients who received a score  $\leq 2$  (the individual is not able to ambulate independently) as severely impaired, whereas those who received a FAC score  $\geq 3$  (the individual is able to ambulate with verbal supervision, without physical contact) as moderately impaired (54). The results showed that the most widely used outcome measures were administered in every phase of disease. Therefore, walking independence, velocity, endurance, balance, mobility and muscle strength, are important components of walking that should be considered from early to late phases of disease to maximize gait recovery (Table VI).

#### *The proposal of battery of tests according to ICF domains and 3-Dimensional Model*

The development of a standardized protocol would permit comparison between different studies, allowing the best rehabilitation approaches to be identified. For example, a Cochrane review showed interesting results emerging from the analysis of effects of robot-assisted therapy to improve gait function. However, the authors could not perform a comparison because of the difference in outcome measures used (4).

Sivan and collaborators (7) performed a similar review, identifying the scales used during robot-assisted upper limb rehabilitation trials in stroke patients. They did not arrive to a shared consensus about the clinical outcome examinations; however, they concluded that the ICF is an appropriate framework to use when choosing an outcome measure. Our results are confirmed by existing literature in neurological rehabilitation of patients with stroke (86).

The choice of the most appropriate clinical scales could be improved, taking into account the following items (7): (i) ICF model to identify the main domains of outcome measures; (ii) analysis of essential psychometric properties (reliability, valid-

ity, responsiveness and sensibility), along with MCID and levels of measurement; (iii) identification of the aim of the measurement (ADL, impairment); (iv) distinction of the different clinical histories of stroke and severity and, subsequently, choice of the optimal scale; (v) nature of the study (effectiveness or efficacy); and (vi) modality of test administration (e.g. interview, questionnaire, phone, or self-report). Future studies should also consider the recovery processes mentioned previously (12).

With this in mind, a specific protocol based on the ICF domain and on a 3-Dimensional Model could be proposed in order to evaluate ERAGTT effects on walking in the clinical setting (10) (Table VII). It is important to note, that this proposal is the result of this extensive review of the literature and it is aimed at satisfying discriminative, evaluative and predictive purposes (Table VII). According to this proposal, the examiner may be guided when choosing the most appropriate scales regarding both the type of measurement (discriminative, evaluative or predictive) and the ICF domain. For instance, the MI, MAS, FAC, 10MWT and 6MinWT could be chosen for discriminative measurements of patients features with reference to body function and structure, and activity domain, respectively. In contrast, if an assessor desires to predict a specific ability that the patient may be able to perform after treatment, the RMI and PASS scales may be used (Table VII).

It is important to note that the BBS could be replaced by the Postural Assessment Stroke Scale (PASS) (87). Furthermore, the MAS and Stroke Impact Scale (SIS) (88) could be considered for inclusion, to evaluate body function and structure, and participation respectively.

Finally, personal and environmental factors could be included in a specific section ("contextual factors" in Table VII) in order to evaluate the patients' and caregivers' impressions of ERAGTT. Both patient and caregiver perception have an important influence on any intervention in rehabilitation, and especially when robot-assisted training is performed. As a whole, the time required to administer this proposed protocol is approximately 56 min (i.e. PASS 10 min; MAS 1 min; SIS 9 min), hence a cost-effective and quick tool.

With respect to the importance of evaluating gait impairment, as well as gait improvements, from a qualitative point of view, instrumental analysis, such as EMG, should be associated with this clinical evaluation protocol. This is particularly relevant

when a distinction between recovery of function and compensation needs to be clarified. However, a distinction between clinical and research settings should be also considered. Specific information or analysis methods may be predominantly suitable or relevant in a research setting instead of a clinical setting.

The main limitation of this review is that robotic orthotic devices were excluded. Secondly, this review attempts to be as comprehensive as possible, but it is likely that some articles were missed. Thirdly, it is possible that other outcome measures, more accurate than those found in the ERAGTT, could be suitable.

In conclusion, we propose a strategy to support researchers and clinicians in the selection of outcome measures in order to evaluate the effects of robotic devices for gait rehabilitation. We believe that a shared evaluation protocol based on ICF domains may provide information to detect changes in the basic components of walking and patient's involvement in real-life situations.

Finally, the selection of common outcome measures could implement research in this important field of rehabilitation by promoting clinical trials and multicentre studies. Future investigations should take into account these considerations, in order to achieve homogeneity among clinical studies and thus allow their results to be compared.

## REFERENCES

1. Feigin VL, Lawes CM, Bennett DA, Anderson CS. Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *Lancet Neurol* 2003; 2: 43–53.
2. Stroke prevention and educational awareness diffusion (SPREAD). The Italian guidelines for stroke prevention and treatment. Milano: Ed. Hyperphar Group; 2012.
3. Patel MD, Tilling K, Lawrence E, Rudd AG, Wolfe CD, McKeivitt C. Relationships between long-term stroke disability, handicap and health-related quality of life. *Age Ageing* 2006; 35: 273–279.
4. Mehrholz J, Werner C, Kugler J, Pohl M. Electromechanical-assisted training for walking after stroke. *Cochrane Database Syst Rev* 2007; CD006185.
5. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet* 2011; 14: 1693–1702.
6. Mazzoleni S, Dario P, Carrozza M. C, Guglielmelli E. Application of robotic and mechatronic systems to neurorehabilitation. In: Annalisa Milella Donato Di Paola, Grazia Cicirelli, editors. *Mechatronic Systems Applications*. Rijeka: Intech; 2010, p. 99–116.

Table VII. Proposed battery of tests according to the International Classification of Functioning, Disability and Health (ICF) domains and 3-Dimensional Model (10). The tests are listed according to the measurement aim, as discriminative, evaluative or predictive. The discriminative scales can be used to divide the patients into homogeneous groups for experimental design. The evaluative scales can be used to evaluate the effects of treatment between the beginning and end of therapy. The predictive scales can be used to predict a specific ability the patient will be able to perform

Type of measurement	Level of assessment (ICF) and domains of assessment			
	Body function and structures	Activity	Participation, health-related and quality of life	Contextual factors Environmental factors    Personal factors
Discriminative	MI, MAS	FAC, 10MWT, 6MinWT,		
Evaluative	MI, MAS	FAC, 10MWT, 6MinWT, RMI PASS	SIS	Patient and carer impression
Predictive	MI	FAC, 10MWT, 6MinWT, RMI, PASS	SIS	

SIS: Stroke Impact Scale; PASS: Postural Assessment Stroke Scale. For other abbreviations, see Table II.



7. Sivan M, O'Connor RJ, Makower S, Levesley M, Bhakta B. Systematic review of outcome measures used in the evaluation of robot-assisted upper limb exercise in stroke. *J Rehabil Med* 2011; 43: 181–189.
8. Mehrholz J, Wagner K, Rutte K, Meissner D, Pohl M. Predictive validity and responsiveness of the functional ambulation category in hemiparetic patients after stroke. *Arch Phys Med Rehabil* 2007; 88: 1314–1319.
9. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J. Issues for selection of outcome measures in stroke rehabilitation: ICF Body Functions. *Disabil Rehabil* 2005; 27: 191–207.
10. Barak S, Duncan PW. Issues in selecting outcome measures to assess functional recovery after stroke. *NeuroRx* 2006; 3: 505–524.
11. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: a systematic review. *Lancet Neurol* 2009; 8: 741–754.
12. Levin MF, Kleim JA, Wolf SL. What do motor “recovery” and “compensation” mean in patients following stroke? *Neurorehabil Neural Repair* 2009; 23: 313–319.
13. Lemberg I, Kirchberger I, Stucki G, Cieza A. The ICF Core Set for stroke from the perspective of physicians: a worldwide validation study using the Delphi technique. *Eur J Phys Rehabil Med* 2010; 46: 377–388.
14. Rehabilitation Measures Database. The rehabilitation clinician's place to find the best instruments to screen patients and monitor their progress: developed by Rehabilitation Institute of Chicago, Center for Rehabilitation Outcomes Research, Northwestern University, Feinberg School of Medicine Department of Medical Social Sciences Informatics group 2010. Available from: <http://www.rehabmeasures.org/default.aspx>.
15. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF activity. *Disabil Rehabil* 2005; 27: 315–340.
16. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF Participation. *Disabil Rehabil* 2005; 27: 507–528.
17. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 2000; 81: S15–S20.
18. Kosak M, Smith T. Comparison of the 2-, 6-, and 12-minute walk tests in patients with stroke. *J Rehabil Res Dev* 2005; 42: 103–107.
19. Fisher S, Lucas L, Thrasher TA. Robot-assisted gait training for patients with hemiparesis due to stroke. *Top Stroke Rehabil* 2011; 18: 269–276.
20. Tyson S, Connell L. The psychometric properties and clinical utility of measures of walking and mobility in neurological conditions: a systematic review. *Clin Rehabil* 2009; 23: 1018–1033.
21. Fulk GD, Echternach JL, Nof L, O'Sullivan S. Clinometric properties of the six-minute walk test in individuals undergoing rehabilitation poststroke. *Physiother Theory Pract* 2008; 24: 195–204.
22. Maeda A, Yuasa T, Nakamura K, Higuchi S, Motohashi Y. Physical performance tests after stroke: reliability and validity. *Am J Phys Med Rehabil* 2000; 79: 519–525.
23. Pandyan AD, Johnson GR, Price CI, Curless RH, Barnes MP, Rodgers H. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. *Clin Rehabil* 1999; 13: 373–383.
24. Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys Ther* 2008; 88: 559–566.
25. Wade DT, Hewer RL. Functional abilities after stroke: measurement, natural history and prognosis. *J Neurol Neurosurg Psychiatry* 1987; 50: 177–182.
26. Kurth T, Gaziano JM, Rexrode KM, Kase CS, Cook NR, Manson JE, et al. Prospective study of body mass index and risk of stroke in apparently healthy women. *Circulation* 2005; 111: 1992–1998.
27. Shinar D, Gross CR, Price TR, Banko M, Bolduc PL, Robinson RG. Screening for depression in stroke patients: the reliability and validity of the Center for Epidemiologic Studies Depression Scale. *Stroke* 1986; 17: 241–245.
28. Côté R, Battista RN, Wolfson C, Boucher J, Adam J, Hachinski V. The Canadian Neurological Scale: validation and reliability assessment. *Neurology* 1989; 39: 638–643.
29. Spilg EG, Martin BJ, Mitchell SL, Aitchison TC. A comparison of mobility assessments in a geriatric day hospital. *Clin Rehabil* 2001; 15: 296–300.
30. Hantson L, De Weerd W, De Keyser J, Diener HC, Franke C, Palm R, et al. The European Stroke Scale. *Stroke* 1994; 25: 2215–2219.
31. Schuling J, de Haan R, Limburg M, Groenier KH. The Frenchay Activities Index. Assessment of functional status in stroke patients. *Stroke* 1993; 24: 1173–1177.
32. Ottenbacher KJ, Hsu Y, Granger CV, Fiedler RC. The reliability of the functional independence measure: a quantitative review. *Arch Phys Med Rehabil* 1996; 77: 1226–1232.
33. Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer assessment for testing motor performance in patients following stroke. *Phys Ther* 1993; 73: 447–454.
34. Sayers SP, Jette AM, Haley SM, Heeren TC, Guralnik JM, Fielding RA. Validation of the Late-Life Function and Disability Instrument. *J Am Geriatr Soc* 2004; 52: 1554–1559.
35. Collin C, Wade D. Assessing motor impairment after stroke: a pilot reliability study. *J Neurol Neurosurg Psychiatry* 1990; 53: 576–579.
36. Loewen SC, Anderson BA. Reliability of the Modified Motor Assessment Scale and the Barthel Index. *Phys Ther* 1988; 68: 1077–1081.
37. Dick JP, Guiloff RJ, Stewart A, Blackstock J, Bielawska C, Paul EA, Marsden CD. Mini-mental state examination in neurological patients. *J Neurol Neurosurg Psychiatry* 1984; 47: 496–499.
38. Carr JH, Shepherd RB, Nordholm L, Lynne D. Investigation of a new motor assessment scale for stroke patients. *Phys Ther* 1985; 65: 175–180.
39. Gregson JM, Leathley MJ, Moore AP, Smith TL, Sharma AK, Watkins CL. Reliability of measurements of muscle tone and muscle power in stroke patients. *Age Ageing* 2000; 29: 223–228.
40. Sulter G, Steen C, De Keyser J. Use of the Barthel index and modified Rankin scale in acute stroke trials. *Stroke* 1999; 30: 1538–1541.
41. Goldstein LB, Bartels C, Davis JN. Interrater reliability of the NIH stroke scale. *Arch Neurol* 1989; 46: 660–662.
42. Kurtais Y, Küçükdeveci A, Elhan A, Yilmaz A, Kalli T, Tur BS, et al. Psychometric properties of the Rivermead Motor Assessment: its utility in stroke. *J Rehabil Med* 2009; 41: 1055–1061.
43. Franchignoni F, Tesio L, Benevolo E, Ottonello M. Psychometric properties of the Rivermead Mobility Index in Italian stroke rehabilitation inpatients. *Clin Rehabil* 2003; 17: 273–282.
44. Dunbar CC, Glickman-Weiss EL, Bursztyjn DA, Kurtich M, Quiroz A, Conley P. A submaximal treadmill test for developing target ratings of perceived exertion for outpatient cardiac rehabilitation. *Percept Mot Skills* 1998; 87: 755–759.
45. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988; 19: 604–607.
46. Horgan NF, Finn AM, O'Regan M, Cunningham CJ. A new stroke activity scale—results of a reliability study. *Disabil Rehabil* 2003; 25: 277–285.
47. Hagen S, Bugge C, Alexander H. Psychometric properties of the SF-36 in the early post-stroke phase. *J Adv Nurs* 2003; 44: 461–468.
48. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994; 49: 85–94.
49. Barber M, Fail M, Shields M, Stott DJ, Langhorne P. Validity and reliability of estimating the Scandinavian stroke scale score from medical records. *Cerebrovasc Dis* 2004; 17: 224–227.
50. Hill KD, Bernhardt J, McGann AM, Maltese D, Berkovits D. A new test of dynamic standing balance for stroke patients: reliability, validity and comparison with healthy elderly. *J Physiother Canada* 1996; 48: 257–262.
51. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc* 1986; 34: 119–126.
52. Dias D, Lains J, Pereira A, Nunes R, Caldas J, Amaral C, et al. Can we improve gait skills in chronic hemiplegics? A randomised

- control trial with gait trainer. *Eura Medicophys* 2007; 43: 499–504.
53. Hesse S, Waldner A, Tomelleri C. Innovative gait robot for the repetitive practice of floor walking and stair climbing up and down in stroke patients. *J Neuroeng Rehabil* 2010; 7: 30.
  54. Conesa L, Costa U, Morales E, Edwards DJ, Cortes M, León D, et al. An observational report of intensive robotic and manual gait training in sub-acute stroke. *J Neuroeng Rehabil* 2012; 9: 13.
  55. Morone G, Bragoni M, Iosa M, De Angelis D, Venturiero V, Coiro P, et al. Who may benefit from robotic-assisted gait training? A randomized clinical trial in patients with subacute stroke. *Neurorehabil Neural Repair* 2011; 25: 636–644.
  56. Geroin C, Picelli A, Munari D, Waldner A, Tomelleri C, Smania N. Combined transcranial direct current stimulation and robot-assisted gait training in patients with chronic stroke: a preliminary comparison. *Clin Rehabil* 2011; 25: 537–548.
  57. Peurala SH, Airaksinen O, Huuskonen P, Jakala P, Juhakoski M, Sandell K, et al. Effects of intensive therapy using gait trainer or floor walking exercises early after stroke. *J Rehabil Med* 2009; 41: 166–173.
  58. Maple FW, Tong RKY, Li LSW. A pilot study of randomized clinical controlled trial of gait training in subacute stroke patients with partial body-weight support electromechanical gait trainer and functional electrical stimulation: six-month follow-up. *Stroke* 2008; 39: 154–156.
  59. Pohl M, Werner C, Holzgraefe M, Kroczeck G, Mehrholz J, Wingendorf I, et al. Repetitive locomotor training and physiotherapy improve walking and basic activities of daily living after stroke: a single-blind, randomized multicentre trial (DEutsche GAngrainerStudie, DEGAS). *Clin Rehabil* 2007; 21: 17–27.
  60. Tong RK, Ng MF, Li LS. Effectiveness of gait training using an electromechanical gait trainer, with and without functional electric stimulation, in subacute stroke: a randomized controlled trial. *Arch Phys Med Rehabil* 2006; 87: 1298–1304.
  61. Peurala SH, Tarkka IM, Pitkanen K, Sivenius J. The effectiveness of body weight-supported gait training and floor walking in patients with chronic stroke. *Arch Phys Med Rehabil* 2005; 86: 1557–1564.
  62. Werner C, Von Frankenberg S, Treig T, Konrad M, Hesse S. Treadmill training with partial body weight support and an electromechanical gait trainer for restoration of gait in subacute stroke patients: a randomized crossover study. *Stroke* 2002; 33: 2895–2901.
  63. Hesse S, Werner C, Uhlenbrock D, von Frankenberg S, Bardeleben A, Brandl-Hesse B. An electromechanical gait trainer for restoration of gait in hemiparetic stroke patients: preliminary results. *Neurorehabil Neural Repair* 2001; 15: 39–50.
  64. Hesse S, Uhlenbrock D. A mechanized gait trainer for restoration of gait. *J Rehabil Res Dev* 2000; 37: 701–708.
  65. Hesse S, Uhlenbrock D, Werner C, Bardeleben A. A mechanized gait trainer for restoring gait in nonambulatory subjects. *Arch Phys Med Rehabil* 2000; 81: 1158–1161.
  66. Chang WH, Kim MS, Huh JP, Lee PK, Kim YH. Effects of robot-assisted gait training on cardiopulmonary fitness in subacute stroke patients: a randomized controlled study. *Neurorehabil Neural Repair* 2012; 26: 318–324.
  67. Magagnin V, Bo I, Turiel M, Fornari M, Caiani EG, Porta A. Effects of robot-driven gait orthosis treadmill training on the autonomic response in rehabilitation-responsive stroke and cervical spondylotic myelopathy patients. *Gait Posture* 2010; 32: 199–204.
  68. Lewek MD, Cruz TH, Moore JL, Roth HR, Dhaher YY, Hornby TG. Allowing intralimb kinematic variability during locomotor training poststroke improves kinematic consistency: a subgroup analysis from a randomized clinical trial. *Phys Ther* 2009; 89: 829–839.
  69. Hidler J, Nichols D, Pelliccio M, Brady K, Campbell DD, Kahn JH, Hornby TG. Multicenter randomized clinical trial evaluating the effectiveness of the Lokomat in subacute stroke. *Neurorehabil Neural Repair* 2009; 23: 5–13.
  70. Schwartz I, Sajin A, Fisher I, Neeb M, Shochina M, Katz-Leurer M, et al. The effectiveness of locomotor therapy using robotic-assisted gait training in subacute stroke patients: a randomized controlled trial. *PMR* 2009; 1: 516–523.
  71. Westlake KP, Patten C. Pilot study of Lokomat versus manual-assisted treadmill training for locomotor recovery post-stroke. *J Neuroeng Rehabil* 2009; 6: 18.
  72. Hornby TG, Campbell DD, Kahn JH, Demott T, Moore JL, Roth HR. Enhanced gait-related improvements after therapist versus robotic-assisted locomotor training in subjects with chronic stroke: a randomized controlled study. *Stroke* 2008; 39: 1786–1792.
  73. Mayr A, Kofler M, Quirbach E, Matzak H, Fröhlich K, Saltuari L. Prospective, blinded, randomized crossover study of gait rehabilitation in stroke patients using the Lokomat gait orthosis. *Neurorehabil Neural Repair* 2007; 21: 307–314.
  74. Krewer C, Müller F, Husemann B, Heller S, Quintern J, Koenig E. The influence of different Lokomat walking conditions on the energy expenditure of hemiparetic patients and healthy subjects. *Gait Posture* 2007; 26: 372–377.
  75. Husemann B, Muller F, Krewer C, Heller S, Koenig E. Effects of locomotion training with assistance of a robot-driven gait orthosis in hemiparetic patients after stroke: a randomized controlled pilot study. *Stroke* 2007; 38: 349–354.
  76. Freivogel S, Schmalohr D, Mehrholz J. Improved walking ability and reduced therapeutic stress with an electromechanical gait device. *J Rehabil Med* 2009; 41: 734–739.
  77. Wu M, Landry JM, Yen SC, Schmit BD, Hornby TG, Rafferty M. A novel cable-driven robotic training improves locomotor function in individuals post-stroke. *Conf Proc IEEE Eng Med Biol Soc* 2011; 2011: 8539–8542.
  78. Tilson JK, Sullivan KJ, Cen SY, Rose DK, Koradia CH, Azen SP, et al. Locomotor Experience Applied Post Stroke (LEAPS) Investigative Team. Meaningful gait speed improvement during the first 60 days poststroke: minimal clinically important difference. *Phys Ther* 2010; 90: 196–208.
  79. Cameron D, Bohannon RW. Criterion validity of lower extremity Motricity Index scores. *Clin Rehabil* 2000; 14: 208–211.
  80. Stevens S. On the theory of scales of measurement. *Science* 1946; 103: 677–680.
  81. Smania N, Gambarin M, Paolucci S, Girardi P, Bortolami M, Fiaschi A, Santilli V, et al. Active ankle dorsiflexion and the Mingazzini manoeuvre: two clinical bedside tests related to prognosis of postural transferring, standing and walking ability in patients with stroke. *Eur J Phys Rehabil Med* 2011; 47: 435–440.
  82. Iosa M, Morone G, Fusco A, Pratesi L, Bragoni M, Coiro P, et al. Effects of walking endurance reduction on gait stability in patients with stroke. *Stroke Res Treat* 2012; 2012: 810415.
  83. Smidt N, de Vet HC, Bouter LM, Dekker J, Arendzen JH, de Bie RA, et al. Exercise Therapy Group. Effectiveness of exercise therapy: a best-evidence summary of systematic reviews. *Aust J Physiother* 2005; 51: 71–85.
  84. Sommerfeld DK, Johansson H, Jönsson AL, Murray V, Wessari T, Holmqvist LW, et al. Rivermead mobility index can be used to predict length of stay for elderly persons, 5 days after stroke onset. *J Geriatr Phys Ther* 2011; 34: 64–71.
  85. Küçükdeveci AA, Tennant A, Grimby G, Franchignoni F. Strategies for assessment and outcome measurement in physical and rehabilitation medicine: an educational review. *J Rehabil Med* 2011; 43: 661–672.
  86. Quinn TJ, Dawson J, Walters MR, Lees KR. Functional outcome measures in contemporary stroke trials. *Int J Stroke* 2009; 4: 200–205.
  87. Benaim C, Pérennou DA, Villy J, Rousseaux M, Pelissier JY. Validation of a standardized assessment of postural control in stroke patients: the Postural Assessment Scale for Stroke Patients (PASS). *Stroke* 1999; 30: 1862–1868.
  88. Vellone E, Savini S, Barbato N, Carovillano G, Caramia M, Alvaro R. Quality of life in stroke survivors: first results from the reliability and validity of the Italian version of the Stroke Impact Scale 3.0. *Ann Ig* 2010; 22: 469–479.