IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 23, 2019, accepted February 21, 2019, date of publication February 26, 2019, date of current version March 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2901814

Review of Automated Systems for Upper Limbs Functional Assessment in Neurorehabilitation

EDWIN DANIEL OÑA SIMBAÑA[1], (Member, IEEE), PATRICIA SÁNCHEZ-HERRERA BAEZA[2], ALBERTO JARDÓN HUETE[1], (Senior Member, IEEE), AND CARLOS BALAGUER[1], (Member, IEEE)

[1]Robotics Lab, Department of Systems Engineering and Automation, University Carlos III of Madrid, 28911 Leganés, Spain
[2]Department of Physical Therapy, Occupational Therapy, Rehabilitation and Physical Medicine, Rey Juan Carlos University, 28922 Alcorcón, Spain

Corresponding author: Edwin Daniel Oña Simbaña (eona@ing.uc3m.es)

This work was supported in part by the Spanish Ministry of Economy and Competitiveness via the ROBOHEALTH (DPI2013-47944-C4-1-R) and ROBOESPAS (DPI2017-87562-C2-1-R) Projects, and in part by the RoboCity2030-III-CM project (S2013/MIT-2748) which is funded by the Programas de Actividades I+D Comunidad de Madrid and cofunded by the Structural Funds of the EU.

ABSTRACT Traditionally, the assessment of upper limb (UL) motor function in neurorehabilitation is carried out by clinicians using standard clinical tests for objective evaluation, but which could be influenced by the clinician's subjectivity or expertise. The automation of such traditional outcome measures (tests) is an interesting and emerging field in neurorehabilitation. In this paper, a systematic review of systems focused on automation of traditional tests for assessment of UL motor function used in neurological rehabilitation is presented. A systematic search and review of related articles in the literature were conducted. The chosen works were analyzed according to the automation level, the data acquisition systems, the outcome generation method, and the focus of assessment. Finally, a series of technical requirements, guidelines, and challenges that must be considered when designing and implementing fully-automated systems for upper extremity functional assessment are summarized. This paper advocates the use of automated assessment systems (AAS) to build a rehabilitation framework that is more autonomous and objective.

INDEX TERMS Automatic assessment, biomedical engineering, motor function, neurorehabilitation, rehabilitation robotics, robotics and automation, upper extremity.

TERMINOLOGY

To reduce the ambiguity in the clinical terminology, the definition of the terms that will be used along the text are given as follows.

- **Test or clinical tool:** this is understood as the procedure that the patients must perform in order to assess the functionality of the upper extremities. It encompasses a series of steps and rules for its proper administration. It can be single or multi-item.
- **Item:** the movement or single task that the patient must perform.
- **Outcome measure:** the result of a test that is used to objectively determine the UE function.

I. INTRODUCTION

A particular case of rehabilitation is aimed at treating the problems caused by disorders affecting the nervous and neuromuscular systems, known as neurorehabilitation. In this case, patient needs are usually multi-dimensional, including physical, cognitive, psychological, and medical, and may be very complex. Neurological rehabilitation can be defined as a process or cycle that aims to optimize a person's participation in society and sense of well-being [1]. The starting and ending steps of this rehabilitation cycle are assessment and evaluation, respectively [2]. At the beginning of the rehabilitation process, the assessment step is focused on collecting data about the patient to identify the problems, the causes of functional limitations, and the wishes and goals of the rehabilitation. At the end, the evaluation step refers to assessing the achievement of the goals of the intervention programme [3], [4]. These goals are measured as changes in the functioning or autonomy.

Additionally, a proper evaluation of the therapeutic effectiveness of rehabilitation is also important due to it being a laborious process of expensive interventions [5]. Because of the complexity of neurological diseases, rehabilitation

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng H. Zhu.

processes mostly are long-term treatments. This fact highlights the importance of the assessment step to provide proper economic management in healthcare facilities, and even more importantly, in public institutions. Assessment requires specialized workers and adequate space and material [1]. Thus, factors such as the optimal administration of clinical procedures (optimizing clinicians' time), the appropriate management of resources (workspace and equipment), and proper management of results (patient record) are quite important.

Regarding the procedure's administration, the assessment process is commonly performed by health professionals themselves using standardized clinical tests in order to have objectivity in the evaluation. For the assessment of upper extremity (UE) motor function, such clinical tests are made up of a set of items or procedures that aims to objectively determine the patient's functioning level. However, the evaluation of motor functionality is a manually performed procedure, and it has some drawbacks.

First, current diagnosis of UE motor impairment is based on the observation of select movements (or tasks) by a trained clinical specialist. This estimation aims to be reliable (intra-operator) and objective (inter-operator). However, the nature of visual inspection includes some degree of uncertainty (subjectivity) that may come from a variety of sources (movement variability [6], [7], observer appreciation [8], etc.). Second, neurological rehabilitation is not a process bounded in time. Recovery of motor function in general, and for UE in particular, depends on the characteristics of each individual and the kind of disease (stroke, Parkinson's disease, etc.). Thus, performing several tests to assess longitudinal changes in motor performance can be difficult in terms of patient burden and cost [9], even for healthcare providers.

On this basis, previously mentioned drawbacks could be reduced via automation of traditional assessment tools. Most of the evaluation tests are composed of well-defined exercises or tasks (e.g., point-to-point movements, reaching tasks, object displacement) that are rated by numerical scales, which may be susceptible to automation. By automation, an objective evaluation of the patient's motor functionality could be achieved. Furthermore, the clinician could be provided with more time to assess the results and, based on this, to correct the therapy protocol, modifying the level of difficulty or adding other tasks.

This automation approach in the assessment of motor function has been considered by the research community in recent years. Different methodologies have been used to automatically measure motor function, but the clinical knowledge provided by traditional examination tests has been retained.

In this paper, a systematic review of systems that address the automation of traditional tests for the assessment of UE motor function, used in neurological rehabilitation, is presented. To the best knowledge of the authors, this is the first review to classify the automated methods for upper limbs functional assessment in general, and in terms of motor function in particular. This review presents an analysis of the literature in this field according to the automation level, the employed technology, the focus of assessment, and the method for automatic outcome generation. The remainder of this paper is organized as follows: Section II provides an overview of UE functional assessment and its fundamentals. A description of the traditional tests and procedures is included. In Section III, the results of the literature review are summarized. These results are presented under different scopes. In Section IV, a series of requirements, guidelines, and challenges that must be considered when designing and implementing automated systems for upper limbs functional assessment are presented. Also, the findings and perspectives are discussed. To conclude, some final remarks are presented in Section V.

II. UPPER LIMBS FUNCTIONAL ASSESSMENT: TRADITIONAL METHOD

Overall, the rehabilitation cycle (shown in Figure 1) involves the identification of a person's problems and needs, relating the problems to relevant factors of the person and the environment, defining rehabilitation goals, planning and implementing the measures, and assessing the effects [1]. In a simplified way, it is made up of four steps: assessment, assignment, intervention, and evaluation.

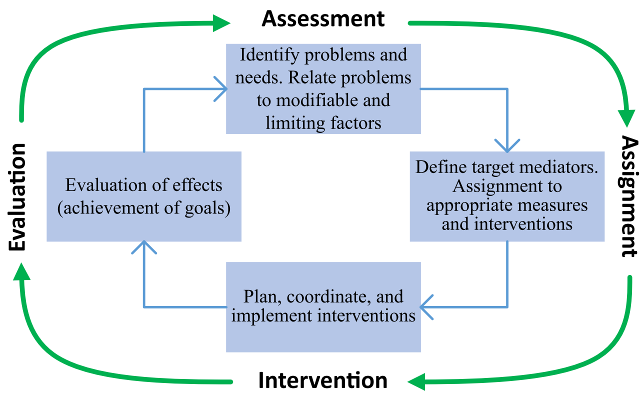

FIGURE 1. The Rehabilitation Cycle [1].

Regarding the assessment stage, functional assessment refers to the determination of a person's ability to perform everyday tasks and requirements of living. Functional assessment is used to establish a baseline, to predict rehabilitation results, and to evaluate therapeutic interventions [10]. Fundamentally, the evaluation process will utilize a number of variables to act as indicators (outcome measures), and these can be compiled to form a clinical assessment to provide a clinically meaningful deduction from the measurement [11].

An outcome measure is the result of a test that is used to objectively determine the functioning level of a patient throughout rehabilitation treatment. Traditionally, outcome measures have focused on the individual's impairment level. However, this provided a limited description of disability. In 2001, the World Health Organization (WHO) adopted the International Classification of Functioning, Disability and Health, commonly known as ICF [12], which provides a common framework for describing the consequences of health conditions and an international standard to describe and measure health and disability. In clinical settings, ICF is used for the evaluation of functional status, goal setting, treatment planning and monitoring, as well as outcome measurements. The ICF model of disability involves three levels: body functions and structure (impairment), activity limitations, and participation.

In the scope of this paper, evaluation of the upper extremities (UE) covers two key factors related to the ICF model: 1) identification of the impairments limiting normal movement, and 2) the initial level of activity limitations and participation restrictions arising from these impairments [13]. For that purpose, standard clinical tests are used to determine the baseline function limitations of a patient at the beginning of treatment. Once treatment has been initiated, the same test(s) can be used to determine progress and treatment efficacy [1], [4]. Nevertheless, the tests for outcome measure gathering are greatly varied with respect to the number, type, and scoring of the tasks used to determine performance levels, their degree of standardization, and their predictive validity [14], [15]. The following section aims to show an overview of the variety of the traditional tests commonly used in neurorehabilitation, including their method of administration and fundamentals for outcome generation.

A. A VARIETY OF AVAILABLE ASSESSMENT TOOLS

Numerous assessment tools are readily available to clinicians to measure disability and function limitations in the neurorehabilitation process. The use of appropriate, valid, and reliable tests can improve the understanding of how disease progresses, the level of structural impairment, and how this impacts on the individual in terms of function and participation [11]. These assessment tools can be categorized according to the functioning levels (ICF model) that we aim to evaluate.

On the one hand, the typical body functions that need to be assessed in the neurological patient are those related to the functions of the joints, muscles, movements, cognitive functions, and sensations. Thus, some constructs of relevance are muscle strength, ranges of movement, attention, memory, and balance. Examples of tests classically encompassed at this level are the Fugl-Meyer Assessment (FMA) of Motor Recovery after Stroke, or the Modified Ashworth Scale (MAS). FMA [16] is a stroke-specific, multi-item, and performance-based impairment index. Test items are scored on the basis of the patient‘s ability to complete the item using a 3-point ordinal scale (0: unable to perform, 1: performs partially, and 2: performs fully). The total possible scale score is 226 for the FMA and 66 for the upper extremities subsection (FMA-UE). Similarly, MAS for measuring spasticity comprises six ordered categories of increasing spasticity that are assigned sequentially in a 5-point scale.

On the other hand, when examining a patient's activities, the therapist will examine not only whether they can do the tasks but also the quality with which the tasks are performed. Example of tools at this level are the Box and Blocks Test (BBT), the Nine-hole Peg Test (NHPG), the Action Research Arm Test (ARAT), or the Wolf Motor Function Test (WMFT). BBT and NHPG are tools for the individual measure of manual dexterity and coordination. For the score, the therapist must manual count the total amount of objects (cubes and pegs for the BBT and NHPG, respectively) transported. WMFT and ARAT quantify upper extremity motor ability through timed and functional tasks (lift objects, reaching, etc). The items are rated on a 6-point scale in the case of WMFT, and a 4-point scale for the ARAT.

Furthermore, participation is a more complex concept than impairments and activities, but it is fundamental to understand the patients and their life and to help with planning treatment. Common scales used are the Canadian Occupational Performance Measure (COMP), the EuroQol Quality of Life Scale (EQ), the Reintegration to Normal Living Index (RNLI), the Stroke Impact Scale (SIS) and the Stroke Specific Quality of Life Scale (SQL). These scales are written questionnaires in which the person has to answer a series of questions that are asked. Detailed descriptions of the features of the tests that are available for functional assessment in neurological rehabilitation are summarized in [15] and [17], according to the ICF model.

Despite the variety of available tests, all of them should accomplish some requirement for clinical acceptance. The outcome measures should evaluate the particular aspect of function that they are reported to assess (validity), and the results should be the same (or similar) regardless of who administers the test or when it is administered (reliability). Additionally, they should actually be able to assess change whatever is being evaluated over time (responsiveness).

B. A PRAGMATIC POINT-OF-VIEW

Classifying the tests within the ICF framework can be difficult and is often controversial [18]. Many of them include items considered an activity within the ICF (e.g., a task performed by an individual), as well as items related to participation (e.g., the societal level of functioning).

In this sense, assessment tools can be also divided into two categories: (1) performance measures, where the clinician rates or times a series of UE actions that are performed by the patient, or (2) self-report measures, where the clinician asks a series of questions about UE actions that are answered verbally by the patient [13]. The traditional tests most commonly used in neurological rehabilitation, according to performance or self-report measures, are listed in Table 1. The tests that fully or partially cover the assessment of UE functionality are marked with * or ** symbols, respectively.

1) PERFORMANCE MEASURES

Performance measures do not measure individual performance but rather analyze the person's process and evolution in activities. These tests shares some characteristics considering two distinctions: some tests analyze the body functions and the others evaluate the activity.

On one side, within the body function domain, we can divide it into three types of tests: the ones that evaluate

TABLE 1. Abbreviations of traditional tests commonly used in neurological rehabilitation, sorted according to Performance and Self-report measures. ICF domains are also indicated (MT: Motor components; CO: Cognitive components; VS: Vital signs; FN: Functional components).

	Abbr.	Test name	ICF
Performance Measures	ARAT	Action Research Arm Test*	MT
	BBT	Box and Block Test*	MT
	BI	Barthel Index	FN
	BIT	Behavioural Inattention Test	CO
	CDT	Clock Drawing Test	CO
	CMSA	Chedoke McMaster Stroke Assessment*	VS
	CNS	Canadian Neurological Scale**	VS
	FAST	Frenchay Aphasia Screening Test	CO
	FIM	Functional Independence Measure	FN
	FMA	Fugl-Meyer Assessment**	MT
	GHQ-28	General Health Questionnaire-28	VS
	MAS	Modified Ashworth Scale**	MT
	MoAS	Motor Assessment Scale**	MT
	MJHFT	Modified Jebsen Hand Function Test*	MT
	MMSE	Mini-Mental State Examination	CO
	MOCA	Montreal Cognitive Assessment	CO
	NHPG	Nine-Hole Peg Test*	MT
	PPT	Purdue Pegboard Test*	MT
	RMA	Rivermead Motor Assessment**	FN
	RPS	Reaching Performance Scale*	FN
	UPDRS	Unified Parkinson's Disease Rating Scale**	MT
	WMFT	Wolf Motor Function Test*	MT
Self-Report Measures	BDI	Beck Depression Inventory	CO
	COPM	Canadian Occupational Performance Measures	FN
	EQ-5D	EuroQol Quality of Life Scale	VS
	GDS	Geriatric Depression Scale	CO
	HADS	Hospital Anxiety and Depression Scale	CO
	LHS	London Handicap Scale	FN
	NHP	Nottingham Health Profile	VS
	RNLI	Reintegration to Normal Living Index	FN
	SA-SIP30	Stroke-Adapted Sickness Impact Profile	VS
	SIS	Stroke Impact Scale*	VS

*Focused on upper extremity (UE); **Includes UE section.

cognitive components, the ones that assess motor components, and the ones that assess clinical states or vital signs.

In the tests that evaluate the cognitive components, a final score and a cut-off point are obtained with which we can compare normality. Most of them are a battery of questions in which the examiner asks the person to perform a series of tasks. In some tests, there may be a time limit, that is to say that the person has to do the activity within a specific time standardized by the test itself; in others, it is not necessary. The type of instruction used in these tests is verbal, that is, the therapist explains what he/she has to do in each activity.

The tests that evaluate the motor components are standardized procedures in which a final score and a cut-off point are obtained with which to compare normality. They are based on a set of items in which the examiner explores different movements of the person. That is, the therapist asks the patient to perform different movements, motor tasks, or adopt different positions with his body to explore whether the person is able or not to perform them. These tests can be controlled by time, as is the case for the BBT in which the patient is asked to move all the necessary cubes for a minute, or they can be simply observational by the therapist in which the action is analyzed and the score is written down. The type of instruction that is used in these tests is verbal in most cases (the therapist explains to the person what he/she has to do).

The scales that assess vital signs are those in which a series of examinations are performed by the examiner to make a clinical judgment, usually derived from the patient's symptoms. They do not have a time limit, and it is the examiner who, based on the evaluations and questions asked in the tests, rates the different scores.

On the other side, the tests in the activity domain are of the so-called functional type, where the individuals are asked to perform different activities or answer questions about how they carry out the activities. Most of them have a set of questions with several response options that range from normality to functional impossibility. It is the examiner who asks the person and records the score based on the scale instructions. The most remarkable thing about these scales is that they evaluate the function, and many of them are linked to the performance of the activities of daily life, so they give the perspective of whether the person can become independent in their day-to-day life.

2) SELF-REPORT MEASURES

These self-report measures are usually a series of questionnaires in which the respondents read the question and select a response for themselves without the researcher's interference. Those questionnaires serve to inquire about the feelings or attitudes of the person, being more often used in observational studies. Because of these measures or self-reports are subjective, problems related to validity could occur, since the patient can be confused by reporting less of the severity of the pathology or, on the contrary, increasing it. In these measures, the intervention of the examiner is not mandatory. Thus, the person can take the questionnaire to his/her home and complete it in a quiet way. Usually the instructions come at the beginning, which explain how to complete the questionnaire. Some of these have a final score, but most are based on the subjective perceptions of the patient. Normally, it is estimated mental states and participation in the environment that give the examiner an idea of the patient's emotional status and the degree of autonomy in different daily activities.

In summary, a wide variety of assessment tools are available for the estimation of functional status by clinicians. Despite the variety, not all of them are (fully or partially) focused on the UE functional assessment. However, it can be appreciated that the dynamics of the assessment of UE functioning, in most of the tests, are susceptible to automation (performance of single movements or tasks, instructions given verbally, observation-based ratings). This fact has been considered by the research community in recent years regarding the development of automated assessment systems. Thus, the following section presents the results of a systematic review focused on analyzing automated systems for assessment of UE functional status where the traditional clinical tests are taken as a design reference.

III. LITERATURE REVIEW SUMMARY

In this section, we highlight the particular aspects of the automated systems for functional assessment of upper limbs in neurorehabilitation. This paper does not intend to be a comprehensive analysis of the utility of the automated assessment systems; rather, it aims to compile the information published in peer-reviewed articles. On this basis, different aspects of automated systems, such as focus of measurements, reliability of data provided, approaches for rating, or clinical feasibility can be discussed.

A. SEARCH METHODS

The authors undertook a literature search in August 2018 about the use of automated systems for the assessment of upper limb motor function in neurological rehabilitation, using keywords such as automated, robot, neurological, rehabilitation, upper, limb, extremity, neurorehabilitation, motor, function, and various combinations of these. The databases were Science Direct, PubMed/Medline, and IEEE. Only papers written in English were considered, and the search was extended to the whole database. Studies were included when: 1) systems for assessment of upper limbs motor function (uni- and bilateral) were addressed; 2) systems were based on traditional tests used in neurorehabilitation, including those that only address the outcome automation; 3) the measure to be automated was a performance measure; 4) the automation of at least one test item was included; 5) clinical trials with real patients were conducted.

B. AUTOMATED APPROACHES FOR FUNCTIONAL ASSESSMENT OF UPPER LIMBS

The results of the systematic review of systems that address the automation of traditional assessment tests used in neurological rehabilitation are summarized in Table 2. The chosen studies were allocated different identification numbers (ID) to better explain them throughout the text. Studies were sorted according to the method used for obtaining the test outcome. To the best knowledge of the authors, this is the first review to classify the automated methods for upper limbs functional assessment.

Next, this paper presents an analysis of the literature review results according to the frequency of use in the automation of the traditional tests, the automation level of the chosen systems, the employed technology, the method for automatic outcome generation, and the focus of the assessment.

1) MOST COMMON AUTOMATED TESTS

A total of 24 automated systems were found in this review. It should be noted that all studies were focused on the automation of a performance-based test for assessment of upper limb functionality.

The frequency of use of each test was calculated based on how many times a specific clinical tool appeared in the third column of Table 2 across the total studies ($n = 24$). The FMA test is clearly the most frequent (46% of studies) test that is considered to be automated. The ARAT, BBT, and WMFT tests are in second place with a frequency of use of 12.5% for each one. Finally, the MJHFT, NHPG, RPS, and UPDRS tests were the third most frequently chosen tests (4% for each one) for automation.

Most of the systems were focused on the automation (partially or completely) of a single test. However, some studies have chosen different items from two tests (ID: 12), or they were able to provide the score of more than one test (ID: 22).

The results of this review are consistent with the Santisteban [55] findings, which showed that the FMA is the most commonly used upper limb outcome measure in intervention studies in stroke rehabilitation. The Santisteban [55] study also concluded that the frequency of use of the tests varies widely, between 36% and 1%. Only 15 measures were used in more than 5% of studies. The WMFT, ARAT, and BBT are included in this range.

The above mentioned tests are able to measure several aspects of motor function, and also can provide a clear perspective about the patient's health status. In this sense, it could be reasonable to develop automated systems based on the most frequently used tests, and therefore, the ones most appropriate to measure functionality.

2) LEVEL OF AUTOMATION

As was previously described in Section II, the assessment process involves test administration and the rating of the test's tasks. That is, a system must address both approaches for full automation. In Table 2, studies that have considered automatic administration of the test are marked via the ✓ symbol. Additionally, the percentage of automated items is indicated.

On the one hand, only six of the reviewed studies dealt with the administration of the assessment in an automatic manner. For this purpose, the most common approach is to give the test's instructions to the patients via a Graphical User Interface (GUI). Different channels can be used for giving the instructions, and it depends on the technology used for automation. In all the systems the instructions are given by audio messages that explain and describe the task. However, most of these also include a visual channel, which displays a video recording in the GUI for demonstration of how the movement (or task) must be performed. The video can show a clinician or an avatar performing the movement.

On the other hand, most of the functional assessment tests are not focused on evaluating specific cognitive or motor functions. That way, the tests can be composed of subsections or domains focused on the assessment of a specific extremity (upper and/or lower) and can evaluate different sensorimotor functions. An example is the FMA test [16], which is made up of five domains, and there are 155 items in total: motor functioning (in the upper and lower extremities); sensory functioning (evaluates light touch on two surfaces of the arm and leg, and position sense for eight joints); balance (contains seven items, three seated and four standing); joint range of motion (eight joints); and joint pain. Hence, 33% of the studies in this review were able to automatically evaluate

TABLE 2. Automated assessment systems based on traditional tests for functional evaluation of upper limbs.

ID	Source	Based on	Automation level		User interface	Technology used for automation			Test outcome obtained by	Outcome provided	Focused on
			Adm*	Items		Sensors	Method	Characteristics††			
(1)	Simonsen, D. (2017) [19], [20]	MJHFT	✗	3 items (100%)	None	IMS (Kinect)	Computer vision for detecting the interaction	Accuracy: ✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**DS**: Image segmentation and point cloud analysis to limit the time measurement	Time spent performing tasks†	Manual dexterity
(2)	Cho, S. (2016) [21]	BBT	✗	All items (100%)	Yes: Virtual interface to develop the test	IMS (Kinect V1)	Virtual reality for modelling the environment and detecting interactions	Accuracy: ✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**DS**: Detecting objects in a virtual environment for cube counting. A classifier based on linear discriminant analysis (LDA) for hand poses (grasping and non-grasping) detection	Number of cubes†	Manual dexterity (gross)
(3)	Oña, E.D. (2018) [22], [23]	FMA	✓	6 items FMA-UE (19%)	Yes: Virtual interface to develop the test and to guide the user by means of auditive messages and video demonstration	IMS (Kinect V2)	Virtual reality for modelling the environment and detecting interactions	Accuracy: ✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**DS**: Detecting collisions between the avatar's joints and virtual detectors (colliders) placed along the required movement path	3-points scale†; *Additional*: Movement smoothness	Motor function (Shoulder, Elbow)
(4)	Hsiao, C.P. (2013) [24]	BBT	✗	All items (100%)	None	IMS (Kinect V1)	Computer vision for detecting interactions	Accuracy: ✦✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**DS**: Image segmentation and point cloud analysis for cube counting	Number of cubes†; *Additional*: Hand movements, speeds and locations	Manual dexterity (gross)
(5)	Oña, E.D. (2018) [25], [26]	BBT	✓	All items (100%)	Yes: GUI to guide the user, provide the test instructions, show results and store data	IMS (Kinect V2)	Computer vision for detecting interactions	Accuracy: ✦✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**DS**: Image segmentation and point cloud analysis for cube counting	Number of cubes†; *Additional*: Partial times; Cube colours	Manual dexterity (gross)
(6)	Gagnon, C. (2014) [27], [28]	NHPG	✗	All items (100%)	Yes: GUI to display in a virtual environment the test apparatus	MMS (Phantom Omni haptic sensor and Force sensors)	Virtual reality for modelling the environment + Haptic sensor to measure interactions and give feedback	Accuracy: ✦✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**DS**: Counting of displaced virtual pegs and time measurement	Time taken to perform tasks†; *Additional*: Grasping force, Acceleration zero-crossings, Number of times a peg is dropped, Trajectory error, Mean collision force	Manual dexterity (fine)
(7)	Cruz, V.T. (2014) [29], [30]	WMFT	✗	2 items (12%)	None	IMU (Magnetic, Angular Rate, Gravity [MARG] sensors)	Analysis of biomechanical data	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦✦	**DS**: Time measurement by the identification of velocity changes. **CS**: Analysis of kinematics movement by means of a Decision Tree classifier for FAS scoring	Time spent performing tasks and the Functional Ability Scale (FAS): 6-points scale†	Motor function (Forearm, Elbow, Hand)
(8)	Scano, A. (2018) [31], [32]	RPS	✗	6 items (100%)	None	IMS (Kinect V2)	Analysis of movements (optical motion tracking)	Accuracy: ✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**CS**: Movement classification using a normative database of healthy subjects' performance (reference ranges of biomechanical performance based on kinematical, dynamical, and motor control parameters)	4-points scale†	Motor function (Shoulder, Elbow, Grasping); Trunk displacement
(9)	Lee, S.H. (2018) [33], [34]	FMA	✓	26 items FMA-UE (79%)	Yes: GUI to provide linguistic guidelines and instruction video to the patient	IMS (Kinect V2) and MMS (Force Sensing Resistor [FSR] sensor)	Analysis of movements (optical motion tracking) and force sensing	Accuracy: ✦✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**CS**: Binary logic-based classification of each item to rate the 3-point scale, without learning procedures	3-points scale†; *Additional*: Range of motion, Grip strength, Time of movement execution	Motor function (Shoulder, Elbow, Forearm, Hand); Co-ordination; Grasping force

*Considering the automatic administration of the test (✓ = Yes; ✗= No); †Same metric as the traditional one; ††Given in levels (Low:✦ , Medium:✦✦ , High:✦✦✦)

TABLE 2. ***(Continued.)*** **Automated assessment systems based on traditional tests for functional evaluation of upper limbs.**

ID	Source	Based on	Automation level		User interface	Technology used for automation			Test outcome obtained by	Outcome provided	Focused on
			Adm*	Items		Sensors	Method	Characteristics††			
(10)	Otten, P. (2015) [35]	FMA	✓	25 items FMA-UE (73%)	Yes: GUI to allow the user to choose the test movement, display instructions. Voice command to start data recording and automatic finish movement detection	IMS (Kinect) and IMU (Inertial sensors) and MMS (glove sensor and Force Sensing Resistor [FSR] sensor)	Analysis of movements (optical + wearable motion tracking) and force sensing	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦	**CS**: Using Support Vector Machine SVM or Backpropagation Neural Network BNN algorithm (machine learning) for movement classification	3-points scale†; *Additional*: Limb orientation, Joint angle, Movement smoothness, Amount of movement, Grip strength, Finger flexion and extension	Motor function (Shoulder, Elbow, Forearm, Wrist, Fingers); Coordination; Grasping force
(11)	Panardi, A. (2010) [36], [37]	WMFT	✗	15 items (88%)	None	IMS (Overhead camera) and IMU (Inertial sensor and computer)	Computer vision for detecting interaction + wearable inertial sensor	Accuracy: ✦✦ Portability: ✦✦✦ Adaptability: ✦✦	**DS**: Real time image processing and sensor's flag for measuring task time. **CS**: Signal processing and Naive Bayes classifier to estimate (algorithm is trained with 5 tasks) the FAS score	Time taken to perform tasks† and the Functional Ability Scale (FAS): 6-points scale†	Motor function (Shoulder, Elbow, Forearm); Grasping ability
(12)	Olesh, E. (2014) [38]	FMA and ARAT	✗	10 movements taken from FMA and ARAT	None	IMS (Kinect)	Analysis of movements (optical motion tracking)	Accuracy: ✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**CS**: Algorithm to reconstruct movements and Principal Component Analysis (PCA) for score estimation	3-points scale†; *Additional*: Joint angles	Motor function (Shoulder, Elbow, Wrist)
(13)	Lee, T.K.M. (2016) [39], [40]	ARAT	✗	1 item (5%)	None	IMU (Accelerometers) and MMS (Force Sensing Resistor [FSR] sensors)	Instrumenting the objects used in the test	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦✦✦	**CS**: Feature combination (6 features) of data signals and a Decision Tree classifier for movement classification	4-points scale† incorporating the time limits	Grasping task
(14)	Piro, N. (2016) [41]	UPDRS	✗	1 item of PART III (6%)	Yes: GUI to allow offline tele-assessment by showing an avatar imitating the user movements	IMU (Magnetic, Angular Rate, Gravity [MARG] sensors)	Analysis of movements (wearable motion capture)	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦	**CS**: Feature selection and movement classification using a J48 algorithm trained on 86 data sets	5-points scale†; *Additional*: Speed of movement (hand turnings) and the angle of the rotation reached	Forearm (pronation-supination, rhythm, speed, and amplitude decrement)
(15)	Kim, W.S. (2016) [42]	FMA	✓	13 items FMA-UE (40%)	Yes: GUI to provided instructions by means of video playback and audio guidelines	IMS (Kinect V1)	Analysis of movements (optical motion tracking)	Accuracy: ✦✦ Portability: ✦✦✦ Adaptability: ✦✦✦	**CS**: Principal component analysis (PCA) and artificial neural network (AAN) learning from the saved motion data	3-points scale†; *Additional*: Smoothness	Motor function (Shoulder, Elbow, Forearm supination)
(16)	Carpinella, I. (2014) [43]	ARAT	✗	All items (100%)	None	IMU (Single inertial sensor at wrist level)	Analysis of movements (wearable motion capture)	Accuracy: ✦✦✦ Portability: ✦✦✦ Adaptability: ✦✦	**CS**: Task segmentation and detecting (semi-automated procedure) the instants of onset and termination of each phase. Extraction of quantitative parameters (duration, jerk index) for movement classification	Z-scores related to the mean item duration and jerk index	Grasp, Grip, Pinch, Gross movements
(17)	Patel, S. (2010) [44], [45]	WMFT	✗	Total FAS score (88%)	None	IMU (accelerometers)	Analysis of movements (wearable motion capture)	Accuracy: ✦✦✦ Portability: ✦✦✦ Adaptability: ✦✦	**CS**: Feature selection and Random Forest based algorithm to classify eight reference tasks. **IS**: Linear regression model using 6 tasks to predict the total FAS score	FAS: 6-points scale†	Motor function (Shoulder, Elbow, Hand); Reaching and Manipulation

*Considering the automatic administration of the test (✓ = Yes; ✗= No); †Same metric as the traditional one; ††Given in levels (Low:✦ , Medium:✦✦ , High:✦✦✦)

TABLE 2. ***(Continued.)*** **Automated assessment systems based on traditional tests for functional evaluation of upper limbs.**

ID	Source	Based on	Automation level: Adm*	Automation level: Items	User interface	Technology used for automation: Sensors	Technology used for automation: Method	Technology used for automation: Characteristics††	Test outcome obtained by	Outcome provided	Focused on
(18)	Villán, M.A. (2018) [46], [47]	FMA	✗	5 items FMA-UE (16%)	None	OMS (Infrared BTS-SMART-D system)	Analysis of movements (marker motion tracking)	Accuracy: ✦✦✦ Portability: ✦ Adaptability: ✦	**CS**: Generation of reference healthy kinematic models (HKM) for a later comparison with the users' motions	3-points scale†; *Additional*: Range of motion	Motor function (Shoulder, Elbow)
(19)	Wang, J. (2014) [48]	FMA	✗	4 items FMA-UE (12%)	None	IMU (wireless inertial sensors)	Analysis of movements (wearable Wireless Body Sensor Network [WBSN] for motion capture)	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦✦	**IS**: Support Vector Regression (SVR) model using 14 selected features based on ReliefF-SVR. The total score is estimated by sensor data of a single task	3-points scale†; *Additional*: Velocity, Energy, Angle of movement	Motor function (Shoulder; Elbow)
(20)	Yu, L. (2016) [49]	Short FMA	✗	7 items SFMA (100%)	None: Remote video conferencing software to administer the test at home	IMU (Accelerometers) and MMS (Flex sensors)	Analysis of movements (wearable motion capture)	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦✦	**IS**: A extreme learning machine (ELM) algorithm to map the sensor data to clinical scores (RRelief algorithm was applied to find the optimal features subset)	3-points scale†	Motor function (Shoulder, Elbow, Wrist, Fingers)
(21)	Del Din, S. (2011) [50]	FMA	✗	Total score FMA-UE (motor)	None	IMU (Inertial sensors)	Analysis of movements (wearable motion capture)	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦✦	**IS**: Random Forest algorithm fed with features derived from wearable sensor data recorded during the performance of a single item of WMFT (lifting a can)	3-points scale†; *Additional*: Energy, velocity, Jerk metric	Motor function (Shoulder, Elbow, Wrist, Hand)
(22)	Bosecker, C. (2010) [51]	FMA, MSS, MP, MAS	✗	Total score FMA	None	MMS (MIT-Manus robot)	End-effector robot for detecting and measuring interactions	Accuracy: ✦✦✦ Portability: ✦ Adaptability: ✦✦✦	**IS**: Linear regression models to estimate clinical scores from the robot-derived metrics from the unconstrained reaching movements (point-to-point) and the circle drawing task	Same as traditional scale for each test†; *Additional*: Shoulder strength	Motor function (Shoulder, Forearm)
(23)	Julianjatsono, R. (2017) [52]	FMA	✗	6 items FMA-UE (18%)	None	IMS (Kinect V2) and IMU + MMS (Glove with Inertial sensor, Force Sensing Resistor, and Flex sensor)	Analysis of movements (wearable motion capture) + sensorized wearable glove	Accuracy: ✦✦✦ Portability: ✦✦ Adaptability: ✦✦	**IS**: Several regression algorithms were implemented to generate high-resolution scores and to classify the movements	High-resolution score (14 fractional digits)	Motor function (shoulder, Elbow, Wrist)
(24)	Prochazka, A. (2015) [53], [54]	N/A	✓	RAHFT tasks	Yes: GUI and software that introduces each component of the test with a three-dimensional animation accompanied by an audio recording	MMS (ReJoyce station)	End-effector robot for detecting and measuring interaction	Accuracy: ✦✦✦ Portability: ✦ Adaptability: ✦✦	**DS**: Subjects perform a variety of movement tasks while playing computer games	Percentage of maximal displacement and Time to completion	Motor function (shoulder, Elbow, Wrist), Grasping ability and force

*Considering the automatic administration of the test (✓ = Yes; ✗= No); †Same metric as the traditional one; ††Given in levels (Low:✦ , Medium:✦✦ , High:✦✦✦)

all of the items (IDs: 7, 8, 16) or the tasks (IDs: 1, 2, 4, 5, 6) of the reference test. Seventeen percent of studies had greater than 70% automated items (IDs: 9, 10, 11, 17).

Regarding the automation level of automated systems, all the test items and their automatic administration must be considered for fully automation. Therefore, only three studies (IDs: 5, 9, 10) can be treated as fully-automatic (or almost) systems.

3) TECHNOLOGIES UTILIZED FOR AUTOMATION

It can be seen that all of the reviewed studies can be considered as performance-based measures (see previous Section II-B.1). Therefore, the test score is based on how the movements are performed by the patient or how long they take.

Consequently, capturing the patient's movements is essential for rating test items. This process of recording human movement is referred as human motion capture, and the systems designed for that purpose as known as MoCap (motion capture) systems.

Several MoCap technologies can be integrated for the automation of traditional tests. In general, five working principles can be distinguished in human motion capture [56]: optoelectronic measurement systems (OMSs), electromagnetic measurement systems (EMSs), image processing systems (IMSs), ultrasonic localization systems (UMSs) and inertial sensory systems (IMUs). Additionally, a different family of techniques can be included: mechanical measurement systems (MMSs) [57]. By means of direct physical interaction, they are able to detect motion (end-effector robot or flex sensors) or can even measure ranges of motion (exoskeletons with angular encoders) of the user.

Four such techniques (OMS, IMS, IMU, MMS) have been identified as commonly used in automatic assessment approaches. The frequency of use (number of studies/total studies) of each technique and different combinations are presented in Figure 2. Details of the employed sensors in each study are included in Table 2.

A total of 33.3% of studies only used vision-based sensors (IMS) for movement tracking, 25% of studies only used inertial sensors (IMU), 12% of studies only used mechanical systems (MMS), and 4.16% only used optoelectronics systems (OMS). The most common combinations were IMS + MMS, IMU + MMS, IMS + IMU, and IMS + IMU + MMS, with frequencies of use in studies of 8.3%, 8.3%, 4.16%, and 4.16%, respectively.

Nevertheless, some clinical tests not only consider the capability of properly performing a task, but also another related feature, such as strength, which is directly related to the ability to interact with the environment. This is the case for the FMA test, which includes an item to specifically gather a resistance measurement when tugging at an object that the user holds. Automatic systems can objectively measure the exerted force during task performance using force sensors (IDs: 9, 10). Other studies (IDs: 6, 22) provide the force

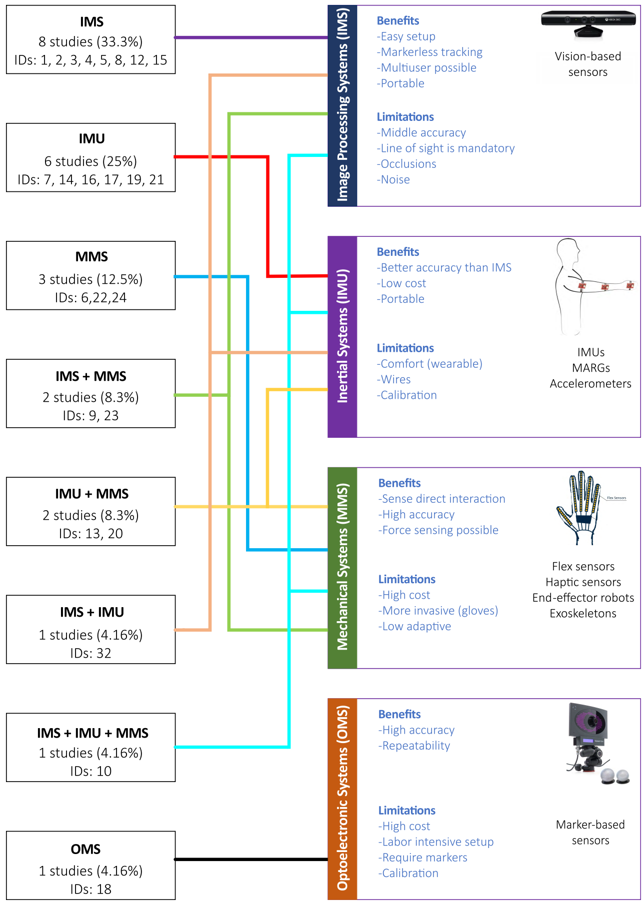

FIGURE 2. Frequency of use and main features of motion capture technologies in automatic assessment approaches.

measurement as an additional outcome, even though it is not considered in the traditional test.

Moreover, the methodologies for using sensors in test automation are varied. Systems that monitor the user-environment interactions by means of computer vision techniques (IDs:1, 4, 5, 11) were identified. Another approach is to adapt the environment (IDs: 2, 3, 6) or the tools (IDs: 13) to sense the user interaction. In addition, a novel approach is to use an end-effector robot that, via its embedded sensors, is able to measure the interaction (IDs: 22, 24). A more extended approach is the analysis of movements based on the registered performance-based data that is considered by the remaining systems.

These methodologies offer some relevant features in regard to automation, such as accuracy, portability, and adaptability to the user's body complexion. First, the accuracy in the data acquisition is high in most of the approaches. Optical sensors allow non-intrusive motion capture. However their accuracy depends on the lighting conditions and a line of sight is required. Wearable sensors give better accuracy at the expense of patient comfort. However, nonetheless, they are not incompatible given an intermediate solution. Furthermore, systems based on these kinds of sensors are portable, being adequate for use outside of clinical settings. This is a drawback for the more accurate motion capture systems like OMS-based or robot-based systems. Finally, due to the nature of neurological disorders, the degree of motor limitations is wide. The target population can vary from elders and children to wheelchair-bound persons. This condition requires that systems can easily adapt to the patient's characteristics. Thus, adaptability is a clear requirement of automated systems to increase their usability. Systems based on non-intrusive sensors seem adequate for fitting to the physical condition of the patients with an easy setup. Proper combination of the sensors and the automation method is a big challenge to obtain the best solution in terms of accuracy, portability, and adaptability.

4) OUTCOME GENERATION

The most relevant advantage of automated systems is the possibility of generating objective outcomes. The general process for automatic generation of outcome measurement is depicted in Figure 3. Different procedures can be applied to obtain a measurement of function based on kinematic data of patients. As a starting point, the goal is to automatically achieve the traditional score. However, novel scores or extended versions of traditional ones can be achieved, considering the richer information that it is obtained by automatic data acquisition systems.

There are two common steps prior to the scoring process, that is, the data acquisition process and signal processing. Different indicators of user performance can be gathered according to the data acquisition method (e.g., IMS: trajectories, range of motion; IMU: kinematic data; MMS: exerted force, etc.). However, these native measurements can be affected by noise. Therefore, a signal processing step is almost mandatory for proper data analysis. Then, different features can be extracted from the enhanced dataset. Such features can feed algorithms for outcome generation.

Regarding the scoring process, three approaches have been identified for automatic generation of clinical outcomes: Direct Scoring (DS), Classification-based Scoring (CS), and Indirect Scoring (IS).

Direct Scoring (DS) systems are those whose outcome is obtained by sensing and analyzing the interactions between the user and the environment. The output is directly calculated from the measurements, and it does not require a trained dataset. A clear example of this approach is when the outcome is given by a measurable variable, such as a time period (IDs: 1, 6). Additionally, countable variables, such as the number of blocks (IDs: 2, 4, 5) can be obtained by direct scoring. Virtual reality can be also useful for the detection and measurement of user-system interactions, providing scores, such as the number of displaced objects (IDs: 2, 6) or a performance-based impairment index (ID: 3).

Classification-based Scoring (CS) is denoted for those systems based on algorithms (with or without learning procedures) that best map input features to an output variable. In this case, a specific dataset is used as a reference for rating

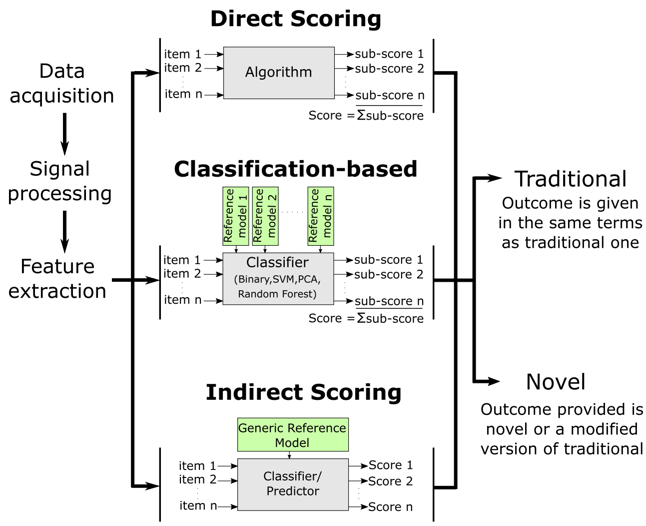

FIGURE 3. Methodologies for automatic outcome generation.

the movements of the test. That is, during the evaluation procedure, each movement is compared with its reference model (features) and it is mapped to determine the best fit. A reference model for each movement (or task) of the test is used. Classic classification algorithms, such as Decision Tree (IDs: 7, 13), Support Vector Machine (ID: 10), Random Forest (ID: 17), and Neural Network (IDs: 10, 15) can be employed. However, in-house-developed algorithms (IDs: 8, 16, 18) were also identified.

Indirect Scoring (IS) is denoted for those systems that use a single reference model for the item rating procedure. Note that a generic/comprehensive reference model is used instead of a model for each movement or task. In ID:19 study, this approach was applied for the prediction of the total score of the FMA test, using sensor data of a single task. However, in this study it also was demonstrated that the prediction performance of single task models was enhanced by building a comprehensive model. In ID:20 study, seven weak regression models for each exercise were established first and then combined to build a comprehensive quantitative FMA (short version) assessment model.

One step beyond, indirect scoring systems are able to estimate another related outcome score. That means, the generic reference model can be used for the prediction of the score of other related outcome measures, without the need to administer the specific tests to measure them. Outcome prediction is based on comparative studies of metrics that are different among them, but keep some correlation. An example is ID:21 study, where the FMA score was estimated using a reference model built based on the data recorded during the performance of a single item (lifting a can) of WMFT. In ID:22 study, different clinical scores (FMA, MSS, MP, and MAS) were estimated from unconstrained reaching movements (point-to-point) and a circle drawing task, using an end-effector robot.

5) EXTENDED OUTCOMES

On the one hand, the main outcome provided by all of the automated systems is the traditional score. In Table 2, the traditional outcomes are labelled by the † symbol. It must be highlighted that, the generated results are more objective than the ones obtained by the clinician observation due to reduced inter-operator variability.

On the other hand, due to the nature of the sensorized systems, additional information about the user performance is directly gathered. In some cases, such extra data can be used for the generation of a modified outcome that gives a better description of impairment than the basic outcome (IDs: 16, 23). Besides, even novel measurements that do not depend on human judgment can be achieved, such as in the method proposed in [53] (ID: 24), and could be an automated alternative to the ARAT or FMA. However, the main drawback of novel outcome measures is the need for clinical validation, as opposed to traditional outcome measures that are already well accepted and widely used by clinicians.

6) FOCUS OF REHABILITATION METHODS

Neurological assessment includes the exploration of cognitive function, language and speech, motor function, reflexes, and sensitive exploration. It can be seen that the automatic systems summarized in this review are based on outcome measures mainly focused on motor function assessment.

The reaching and grasping ability are the motor functions most commonly evaluated by automated systems. The assessment procedure, in more detail, involves the tracking of various joints in order to assess representative motor capabilities such as range of motion, coordination, grasping force, or fine manual dexterity.

IV. TOWARDS AUTOMATED ASSESSMENT SYSTEMS

According to the WHO, the rehabilitation cycle, in a simplified way, is made up of four steps: assessment, assignment, intervention, and evaluation [1]. This rehabilitation cycle, shown in the previous Figure 1, is being transformed into a more automated cycle [58], as shown in Figure 4. This transformation adds more detail but does not alter the rehabilitation cycle, thus maintaining the philosophy centered on the user.

In the past few decades, robotics research has been mainly focused on developing systems in the field of rehabilitation as interventions (systems for recovery/support/training of motor function) [58]. A low percentage of such systems address the assessment stage using the metrics that are obtained during therapy development. Nevertheless, it is important to distinguish that most of the assessments performed by robotic rehabilitation systems (RRS) are not functional assessments, conventionally carried out at baseline and follow-up stages of treatment using standard outcome measures. On the contrary, this type of assessment serves as a method of "rapid evaluation" to inform the therapist about the treatment evolution, by comparing biomechanical data among rehabilitative sessions. However, these outcomes (in terms of trajectories, kinematic data, etc.), despite being indicators of the patient's performance, are nonclinical metrics requiring comparative studies or clinical trials to be validated.

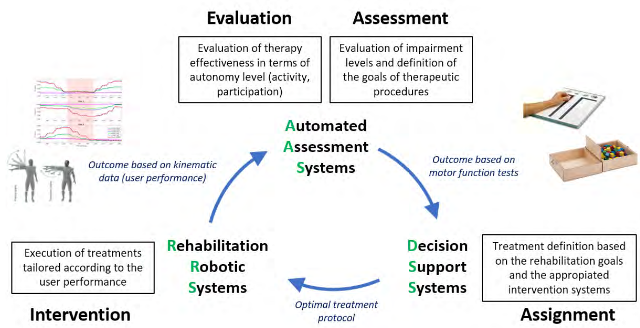

FIGURE 4. The Automated Rehabilitation Cycle [58].

On the one hand, one of the biggest problems with evaluation using traditional tests is the time taken by the therapist to administer them (e.g., FMA [59]). In this way, the report provided by RRS as a rapid evaluation method may be useful. However, the need for clinical validation of results is a drawback. For that reason, the development of automated methods based on traditional assessment scales, that already are clinically validated, widely used, and well-known by specialists in rehabilitation, is certainly desirable. As a result of using automated assessment methods based on traditional tests, the time it takes clinicians to get results will decrease and additional validation will not be required.

On the other hand, the major concern during the evaluation procedure is reducing the subjectivity in assessment. Procedures based on clinician observation could be affected by inter-operator errors. That is, the rating of the same impairment can vary among different clinical professionals. Automation could contribute to reducing inter-operator errors, and could even generate extended performance-based metrics.

In this way, as revealed by the review presented in this paper, the use of traditional clinical tests as a reference for the design of automatic assessment systems (AAS) is a feasible approach. However, there is still room for improvements. Most of the systems are mainly focused on automatic outcome generation. Nevertheless, the assessment process also involves human factors in the test administration that have not yet been completely solved. In the following section, the main challenges, technical requirements, and guidelines that must be considered when designing and implementing the AAS for upper limbs are discussed in order to obtain fully-automated assessment systems.

A. CHALLENGES FOR FULL AUTOMATION

Figure 5 depicts the three main aspects that were identified as mandatory for considering assessment systems to be fully-automated: administration, data acquisition, and rating.

The three components are strongly linked, and they depend

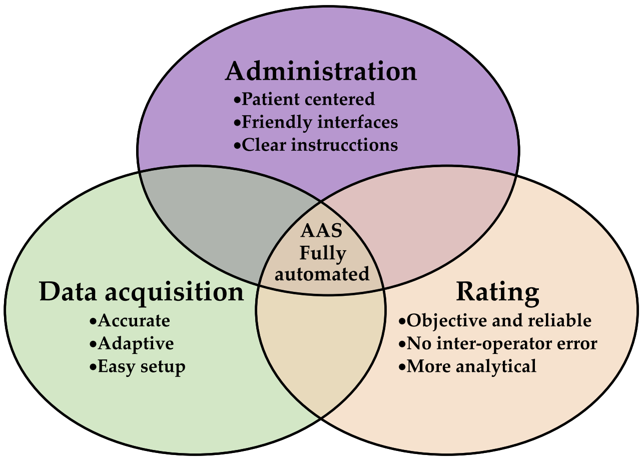

FIGURE 5. Basis for fully-automated assessment systems (AAS).

upon each other for adequate performance. Namely, good data acquisition not only depends on the reliability of the sensors, but also the method of administering the test. An incorrect way of giving the instructions to the patient could lead to incorrect movement execution, caused by trouble understanding the instructions instead of a real impairment. This incorrect data registration will produce an incorrect item rating and thereby an incorrect assessment.

Therefore, proper integration of these aspects could lead to fully-automated assessment systems, based on the combination of clinical knowledge provided by traditional examination tests with the more refined capabilities of biomechanical capture systems. Next, the issues related to the automatic processes of administration, data acquisition, and rating are addressed separately.

1) AUTOMATIC DATA ACQUISITION

The body is characterized by a high number of muscles and joints, all of which must be controlled during the execution of coordinated functional movement [60]. Despite the fact that not all variables involved in movement execution can be measured, some variables related with muscles (activation, strength) or joints (angles, trajectories, velocity, etc.) can be measured by state-of-the-art sensors.

On the one hand, one of the major concerns in the assessment process is obtaining accurate outcome measurements [61]. Additionally, the person being monitored would not even notice the existence of the sensing device or procedure. Unobtrusive sensing technologies, which can be implemented in the form of optical motion tracking or small wearables and Internet of Things (IoT) devices may be a good solution. However, there is a difficulty in deriving useful information from low quality signals. Thus, improving the quality of signals must be the focus of research for future development. So, one of the key requirements for AAS is to detect as many movement indicators as possible, accurately and sustainably in various scenarios.

On the other hand, the assessment scenario can vary according to the patients' characteristics (adults, children, stature, or even clothes) and their mobility restrictions (standing or wheelchair). This fact highlight the acquisition systems must be able to easily adapt to changes in the physical characteristics of the patients, allowing a quick setup. A variety of technologies are currently used to track a person's health and wellness status. They include electrodes, optical sensors, strain gauges, and ultrasound devices, each of which has some drawbacks in terms of user experience such as comfort and convenience. In this way, although the challenges with accuracy and robustness must be still improved, markerless motion capture systems are likely to have a stronger impact on AAS development regarding comfort, adaptability, and easy setup.

Data acquisition systems should be selected according to their accuracy, portability, adaptability, and comfort. Thus, the best solution may be obtained by combination of different types of sensors to increase the accuracy. Proper sensor selection will provide clinicians with useful metrics, and increase the speed and repeatability of the analysis by removing subjective components.

2) AUTOMATED ADMINISTRATION

The main concern about the feasibility of automatic test administration is considering the best way to address the human factors. It should be taken into consideration that the assessment is focused on the patient (patient-centered evaluation). In certain stages of the recovery process, especially in the early stages, the role of the therapist is irreplaceable. Consequently, the usability of AAS will have a niche bounded by the level of affectation of the patients. It seems barely feasible to use for individuals with severe impairments.

Furthermore, the level of interaction between a patient and an automatic system would be not comparable with the clinician-patient interaction level in either case. The therapist's role is not limited to evaluation but also to offer some social skills to encourage the patient throughout the test development, or simply to have a dialog. In this regard, advances in interactive systems (virtual or augmented reality based) or social assistive robots (SAR) are promising in order to implement more reliable dialog systems. Currently, the SARs offer enhanced interaction features, intelligence capabilities, and good acceptance by the users. Thus, including a SAR to monitor the assessment process seems a feasible approach to provide a friendly interaction method, even for telerehabilitation.

In any case, independently of the used interaction method, the test instructions must be provided in a clear manner. This issue is highly important for accuracy in data acquisition, and thereby, in the impairment rating. The more communication channels (audio, text, visual) used, the better.

3) AUTO-RATING

There is an important difference in emphasis between clinical assessment and measurement. It can be seen that automatic data acquisition systems make a greater amount of biomechanical data (measurements) available for the therapist. Some variables (e.g., time and strength) can be measured directly. Other variables (e.g., disability, motor function, or quality of life) are measured indirectly by how they manifest. Interpretation of impairment manifestations is carried out by clinicians using the appropriated outcome measures. Therefore, an automatic assessment method must be able to transform the raw data (performance-based variables) into clinical metrics that can be taken as an objective clinical evaluation (impairment indicators). This is the auto-rating process. As identified in this review, current automated systems use three auto-rating methods (Direct, Classification-based, and Indirect) for the automatic generation of objective clinical outcomes. In addition, the use of auto-rating methods could reduce the inter-operator variability towards reliable measurements by nature.

One step beyond, a more analytic rating procedure is possible by using the AAS. On the one hand, data obtained by traditional outcome measures, both item scores and total scores, are ordinal level, which means that the values are rank ordered [61]. Consequently, these measurements are not precise measurements of an individual. For example, the FMA scale has three categories that are ordered in terms of increasing mobility. Since the main goal of AAS is providing the traditional score, the obtained outcome could have the same drawbacks, despite such methods using richer information gathered by accurate data acquisition systems.

In this way, the sensor data gathered by automated systems could enable the generation of metrics with increased resolution. An example is the ID:23 study, which was able to provide a high resolution outcome for the FMA test while maintaining the classical dynamic in the assessment procedure. This could be a research line to be considered by future works in order to better use sensor data.

On the other hand, automatic rating can be tackled by applying several algorithms (Extreme Learning Machines, Principal Components Analysis, Support Vector Machines, etc). Most of them use a reference model (single or comprehensive) for the classification/prediction of clinical outcomes. Using a comprehensive model appears to be adequate since some studies have demonstrated that multi-item measures need only a few carefully chosen items to generate reliable and valid estimations. Besides, AAS not only must be able to detect the evolution of individuals but must also try to identify whether such changes are clinically significant.

However, most of the studies have considered small samples and they are only suitable for group-level comparisons. Therefore, one of the major challenges in obtaining automatic outcome measures that detect clinically significant changes at the level of the individual is building appropriate reference models including large samples and different populations.

In this regard, based on the capability of AAS for the automatic acquisition and storage of biomechanical data, it is possible to build healthy kinematic models that better fit healthy ranges or patterns. To this purpose, it is necessary to produce a joint effort by researchers and practitioners. Currently, Artificial Intelligence (AI) is a mature science, and digital health (eHealth) transformation using the information and communication technologies (ICT) is a stronger trend in healthcare. Benefits of including such technology in auto-rating in particular, and in AAS development in general, is a research line which is yet to be fully discovered.

B. LIMITATIONS

This review is not without its limitations. Our study was limited to the functional assessment of upper extremities in general, and motor function evaluation in particular. Besides, only automatic systems based on traditional tests were considered. However, there are several developments of automatic assessment systems that propose different methodologies than the traditional ones, even for the lower limbs. Thereby, novel clinical outcomes are obtained that need to be validated. Future works could extend the literature analysis to cover novel automatic systems, including their validity, reliability, and responsiveness.

V. CONCLUSIONS

The objective and automatic measurement of rehabilitation outcomes is a new and developing field. In this paper, a total of 24 works focused on the automatic assessment of UE function in neurorehabilitation were reviewed. On this basis, some remarkable findings in understanding the benefits and challenges when developing automated assessment systems (AAS) were identified.

The development of the AAS should be based on the traditional assessment methods, since the traditional scales are still the "golden standard" for measuring outcomes and determine the effectiveness of treatment. The combination of clinical knowledge provided by traditional examination tests with the more refined capabilities of biomechanical sensors can enhance the outcome measures. Consequently, the outcomes provided by the AAS will be objective, reliable, and will generate additional information about the user's performance.

In addition, we found that the automatic outcome generation of the AAS is based on three methods: Direct Scoring (DS), Classification-based Scoring (CS), and Indirect Scoring (IS). With the exception of the DS case, all of the methods need a (single or comprehensive) healthy reference model to compare the tested movements with the normal ones. Thus, an important issue to solve is the creation of a framework by clinicians and researchers to build appropriate healthy reference models.

Automatic administration of the tests must be also considered, not only the automation of the outcome, to develop fully-automated assessment systems. Knowledge of the user is as important as system functionality, since without the user's cooperation and acceptance, the system's functionality may be ineffective.

In conclusion, it is our opinion that the benefits offered by the AAS can enhance the rehabilitation process, and that these kind of systems will become a complementary tool for common clinical practice.

REFERENCES

[1] World Health Organization, Geneva, Switzerland, *Neurological Disorders: Public Health Challenges*. WHO Press. 2006. [Online]. Available: http://www.who.int/mental_health/neurology/neurodiso/en/

[2] World Health Organization and World Bank, Geneva, Switzerland. *World Report on Disability*. WHO Press. 2011. [Online]. Available: http://www.who.int/en/

[3] M. Barnes, ''Principles of neurological rehabilitation,'' *J. Neurol., Neurosurg. Psychiatry*, vol. 74, no. 4, pp. iv3–iv7, Dec. 2003.

[4] D. Wade, ''Principles of neurological rehabilitation,'' in *Brain Diseases Nervous Syst.*, 12th ed., M. Donaghy, Ed. New York, NY, USA: Oxford Univ. Press, 2009, pp. 165–179.

[5] J. J. M. F. van der Putten, J. C. Hobart, J. A. Freeman, and A. J. Thompson, ''Measuring change in disability after inpatient rehabilitation: Comparison of the responsiveness of the barthel index and the functional independence measure,'' *J. Neurol., Neurosurg. Psychiatry*, vol. 66, no. 4, pp. 480–484, 1999.

[6] R. J. van Beers, P. Haggard, and D. M. Wolpert, ''The role of execution noise in movement variability,'' *J. Neurophysiol.*, vol. 91, no. 2, pp. 1050–1063, Feb. 2004.

[7] R. T. Harbourne and N. Stergiou, ''Movement variability and the use of nonlinear tools: Principles to guide physical therapist practice,'' *Phys. Therapy*, vol. 89, no. 3, pp. 267–282, Mar. 2009.

[8] K. Tanaka and H. Yano, ''Errors of visual judgement in precision measurements,'' *Ergonomics*, vol. 27, no. 7, pp. 767–780, Jul. 1984.

[9] N. Toosizadeh, J. Mohler, H. Lei, S. Parvaneh, S. Sherman, and B. Najafi, ''Motor performance assessment in parkinsons disease: Association between objective in-clinic, objective in-home, and subjective/semi-objective measures,'' *PloS one*, vol. 10, no. 4, 2015, Art. no. e0124763.

[10] M. O. Miller-Keane. *Functional Assessment. Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition*. Accessed: Dec. 3, 2018. [Online]. Available: https://medical-dictionary.thefreedictionary.com/functional+assessment

[11] C. Metcalf, J. Adams, J. Burridge, V. Yule, and P. Chappell, ''A review of clinical upper limb assessments within the framework of the WHO ICF,'' *Musculoskeletal Care*, vol. 5, no. 3, pp. 160–173, Sep. 2007.

[12] World Health Organization, Geneva, Switzerland. *International Classification Functioning, Disability Health (ICF)*. WHO Press, 2001. [Online]. Available: http://www.who.int/classifications/icf/en/

[13] C. E. Lang, M. D. Bland, R. R. Bailey, S. Y. Schaefer, and R. L. Birkenmeier, ''Assessment of upper extremity impairment, function, and activity after stroke: Foundations for clinical decision making,'' *J. Hand Therapy*, vol. 26, no. 2, pp. 104–115, Apr. 2013.

[14] R. Haigh *et al.*, ''The use of outcome measures in physical medicine and rehabilitation within europe,'' *J. Rehabil. Med.*, vol. 33, no. 6, pp. 273–278, Nov. 2001.

[15] K. Salter *et al.*, ''Outcome measures in stroke rehabilitation,'' in *Evidence-Based Review of Stroke Rehabilitation*. Toronto, ON, Canada: Canadian Partnership for Stroke Recovery, 2014.

[16] A. R. Fugl-Meyer, L. Jààskö, I. Leyman, S. Olsson, and S. Steglind, ''The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance,'' *Scandin. J. Rehabil. Med.*, vol. 7, no. 1, pp. 13–31, 1975.

[17] V. Schepers, M. Ketelaar, I. van de Port, J. Visser-Meily, and E. Lindeman, ''Comparing contents of functional outcome measures in stroke rehabilitation using the international classification of functioning, disability and health,'' *Disability Rehabil.*, vol. 29, no. 3, pp. 221–230, Jan. 2007.

[18] S. V. Duff, A. Shumway-Cook, and M. Woollacott, ''Clinical management of the patient with reach, grasp and manipulation disorders,'' in *Motor Control: Translating Research into Clinical Practice*, 4th ed. Philadelphia, PA, USA: Lippincott Williams, 2010.

[19] D. Simonsen, E. G. Spaich, and O. K. Andersen, ''Microsoft kinect-based system for automatic evaluation of the modified jebsen test of hand function,'' in *Converging Clinical and Engineering Research on Neurorehabilitation II*. New York, NY, USA: Springer, 2017, pp. 1299–1303.

[20] D. Simonsen, I. F. Nielsen, E. G. Spaich, and O. K. Andersen, ''Design and test of an automated version of the modified Jebsen test of hand function using microsoft kinect,'' *J. Neuroeng. Rehabil.*, vol. 14, no. 1, p. 38, Dec. 2017.

[21] S. Cho, W.-S. Kim, N.-J. Paik, and H. Bang, ''Upper-limb function assessment using VBBTs for stroke patients,'' *IEEE Comput. Graph. Appl.*, vol. 36, no. 1, pp. 70–78, Jan. 2016.

[22] E. D. Oña, C. Balaguer, and A. Jardón, ''Towards a framework for rehabilitation and assessment of upper limb motor function based on serious games,'' in *Proc. IEEE 6th Int. Conf. Serious Games Appl. Health (SeGAH)*, May 2018, pp. 1–7.

[23] E. D. Oña, A. Jardón, E. Monge, F. Molina, R. Cano, and C. Balaguer, ''Towards automated assessment of upper limbs motor function based on fugl-meyer test and virtual environment,'' in *Converging Clinical and Engineering Research on Neurorehabilitation III*, L. Masia, S. Micera, M. Akay, and J. L. Pons, Eds. Cham, Switzerland: Springer, 2019, pp. 297–301.

[24] C.-P. Hsiao, C. Zhao, and E. Y.-L. Do, ''The digital box and block test automating traditional post-stroke rehabilitation assessment,'' in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, Mar. 2013, pp. 360–363.

[25] E. D. Oña, P. Sánchez-Herrera, A. Cuesta-Gómez, S. Martinez, A. Jardón, and C. Balaguer, ''Automatic outcome in manual dexterity assessment using colour segmentation and nearest neighbour classifier,'' *Sensors*, vol. 18, no. 9, p. 2876, Aug. 2018.

[26] E. D. Oña and A. Jardón, and C. Balaguer, ''The automated box and blocks test an autonomous assessment method of gross manual dexterity in stroke rehabilitation,'' in *Towards Autonomous Robotic Systems*, Y. Gao, S. Fallah, Y. Jin, and C. Lekakou, Eds. Cham, Switzerland: Springer, 2017, pp. 101–114.

[27] C. Gagnon *et al.*, ''The virtual peg insertion test as an assessment of upper limb coordination in ARSACS patients: A pilot study,'' *J. Neurological Sci.*, vol. 347, nos. 1–2, pp. 341–344, Dec. 2014.

[28] M.-C. Fluet, O. Lambercy, and R. Gassert, ''Upper limb assessment using a virtual peg insertion test,'' in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Jul. 2011, pp. 1–6.

[29] V. T. Cruz, V. F. Bento, D. D. Ribeiro, and I. Araújo, C. A. Branco, and P. Coutinho, ''A novel system for automatic classification of upper limb motor function after stroke: An exploratory study,'' *Med. Eng. Phys.*, vol. 36, no. 12, pp. 1704–1710, Dec. 2014.

[30] V. F. Bento, V. T. Cruz, D. D. Ribeiro, and J. P. Cunha, ''Towards a movement quantification system capable of automatic evaluation of upper limb motor function after neurological injury,'' in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2011, pp. 5456–5460.

[31] A. Scano, A. Chiavenna, M. Malosio, L. M. Tosatti, and F. Molteni, ''Kinect V2 implementation and testing of the reaching performance scale for motor evaluation of patients with neurological impairment,'' *Med. Eng. Phys.*, vol. 56, pp. 54–58, Jun. 2018.

[32] A. Scano, A. Chiavenna, M. Malosio, and L. M. Tosatti, ''Kinect v2 performance assessment in daily-life gestures: Cohort study on healthy subjects for a reference database for automated instrumental evaluations on neurological patients,'' *Appl. Bionics Biomech.*, vol. 2017, Nov. 2017, Art. no. 8567084.

[33] S. Lee, Y.-S. Lee, and J. Kim, ''Automated evaluation of upper-limb motor function impairment using Fugl-Meyer assessment,'' *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 1, pp. 125–134, Jan. 2018.

[34] S.-H. Lee, M. Song, and J. Kim, ''Towards clinically relevant automatic assessment of upper-limb motor function impairment,'' in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Inform. (BHI)*, Feb. 2016, pp. 148–151.

[35] P. Otten, J. Kim, and S. H. Son, ''A framework to automate assessment of upper-limb motor function impairment: A feasibility study,'' *Sensors*, vol. 15, no. 8, pp. 20097–20114, 2015.

[36] A. Parnandi, E. Wade, and M. Matarić, ''Motor function assessment using wearable inertial sensors,'' in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Sep. 2010, pp. 86–89.

[37] E. Wade, A. R. Parnandi, and M. J. Mataríc, ''Automated administration of the wolf motor function test for post-stroke assessment,'' in *Proc. 4th Int. Conf. Pervasive Comput. Technol. Healthcare*, Mar. 2010, pp. 1–7.

[38] E. V. Olesh, S. Yakovenko, and V. Gritsenko, ''Automated assessment of upper extremity movement impairment due to stroke,'' *PloS one*, vol. 9, no. 8, 2014, Art. no. e104487.

[39] T. K.-M. Lee *et al.*, ''Towards rehabilitative e-health by introducing a new automatic scoring system,'' in *Proc. 18th Int. Conf. e-Health Netw., Appl. Services*, Sep. 2016, pp. 1–6.

[40] T. K.-M. Lee, J. Lim, S. Sanei, and S. Gan, "Advances on singular spectrum analysis of rehabilitative assessment data," *J. Med. Imaging Health Informat.*, vol. 5, no. 2, pp. 350–358, Jun. 2015.
[41] N. E. Piro, L. K. Piro, J. Kassubek, and R. A. Blechschmidt-Trapp, "Analysis and visualization of 3D motion data for UPDRS rating of patients with parkinson's disease," *Sensors*, vol. 16, no. 6, p. 930, Jun. 2016.
[42] W.-S. Kim, S. Cho, D. Baek, H. Bang, and N.-J. Paik, "Upper extremity functional evaluation by fugl-meyer assessment scoring using depth-sensing camera in hemiplegic stroke patients," *PloS one*, vol. 11, no. 7, 2016, Art. no. e0158640.
[43] I. Carpinella, D. Cattaneo, and M. Ferrarin, "Quantitative assessment of upper limb motor function in multiple sclerosis using an instrumented action research arm test," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, p. 67, Dec. 2014.
[44] S. Patel *et al.*, "A novel approach to monitor rehabilitation outcomes in stroke survivors using wearable technology," *Proc. IEEE*, vol. 98, no. 3, pp. 450–461, Mar. 2010.
[45] S. Patel *et al.*, "Tracking motor recovery in stroke survivors undergoing rehabilitation using wearable technology," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Sep. 2010, pp. 6858–6861.
[46] M. A. Villán *et al.*, "Objective motor assessment for personalized rehabilitation of upper extremity in brain injury patients," *NeuroRehabilitation*, vol. 42, no. 4, pp. 429–439, Jun. 2018.
[47] M. A. Villán *et al.*, "A first step for the automation of Fugl-Meyer assessment scale for stroke subjects in upper limb physical neurorehabilitation," in *Proc. ICIMTH*, Jun. 2015, pp. 45–48.
[48] J. Wang, L. Yu, J. Wang, L. Guo, X. Gu, and Q. Fang, "Automated Fugl-Meyer assessment using SVR model," in *Proc. IEEE Int. Symp. Bioelectron. Bioinf.*, Apr. 2014, pp. 1–4.
[49] L. Yu, D. Xiong, L. Guo, and J. Wang, "A remote quantitative Fugl-Meyer assessment framework for stroke patients based on wearable sensor networks," *Comput. Methods Programs Biomed.*, vol. 128, pp. 100–110, May 2016.
[50] S. D. Din, S. Patel, C. Cobelli, and P. Bonato, "Estimating Fugl-Meyer clinical scores in stroke survivors using wearable sensors," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2011, pp. 5839–5842.
[51] C. Bosecker, L. Dipietro, B. Volpe, and H. I. Krebs, "Kinematic robot-based evaluation scales and clinical counterparts to measure upper limb motor performance in patients with chronic stroke," *Neurorehabil. Neural Repair*, vol. 24, no. 1, pp. 62–69, 2010.
[52] R. Julianjatsono, R. Ferdiana, and R. Hartanto, "High-resolution automated Fugl-Meyer assessment using sensor data and regression model," in *Proc. 3rd Int. Conf. Sci. Technol. Comput.*, Jul. 2017, pp. 28–32.
[53] A. Prochazka and J. Kowalczewski, "A fully automated, quantitative test of upper limb function," *J. Motor Behavior*, vol. 47, no. 1, pp. 19–28, Jan. 2015.
[54] J. Kowalczewski, E. Ravid, and A. Prochazka, "Fully-automated test of upper-extremity function," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2011, pp. 7332–7335.
[55] L. Santisteban, M. Térémetz, J.-P. Bleton, J.-C. Baron, M. A. Maier, and P. G. Lindberg, "Upper limb outcome measures used in stroke rehabilitation studies: A systematic literature review," *PloS one*, vol. 11, no. 5, May 2016, Art. no. e0154792.
[56] E. van der Kruk and M. M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review," *Eur. J. Sport Sci.*, vol. 18, no. 6, pp. 806–819, Jul. 2018.
[57] G. Welch and E. Foxlin, "Motion tracking: No silver bullet, but a respectable arsenal," *IEEE Comput. Graph. Appl.*, vol. 22, no. 6, pp. 24–38, Nov. 2002.
[58] E. D. Oña, R. Cano de la Cuerda, and P. Sánchez-Herrera, C. Balaguer, and A. Jardôn, "A review of robotics in neurorehabilitation: Towards an automated process for upper limb," *J. Healthcare Eng.*, vol. 2018, pp. 1–19, 2018, Art. no. 9758939.
[59] D. J. Gladstone, C. J. Danells, and S. E. Black, "The fugl-meyer assessment of motor recovery after stroke: A critical review of its measurement properties," *Neurorehabil. Neural Repair*, vol. 16, no. 3, pp. 232–240, Sep. 2002.
[60] A. Shumway-Cook and M. H. Woollacott, *Motor Control: Translating Reserch into Clinical Pract.*, 4th ed. Philadelphia, PA, USA: Lippincott Williams Wilkins, 2007.
[61] J. C. Hobart, S. J. Cano, J. P. Zajicek, and A. J. Thompson, "Rating scales as outcome measures for clinical trials in neurology: Problems, solutions, and recommendations," *Lancet Neurol.*, vol. 6, no. 12, pp. 1094–1105, Dec. 2007.

EDWIN DANIEL OÑA SIMBAÑA (M'18) was born in Quito, Ecuador. He received the B.Sc. degree in electronics engineering and the M.Sc. degree in advanced electronic systems from the University Carlos III of Madrid (UC3M), in 2011 and 2013, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering, electronics, and automation.

From 2013 to 2015, he was a Research Assistant with the Group of Power Electronics Systems, UC3M, focused on analyzing and modelling of a modular ISOP Full Bridge-based converter with input filter, applied to transportation systems. Since 2015, he has been a Research Assistant with the Robotics Lab, UC3M. His research is focused on automatic methods for rehabilitation and assessment of motor function in neurological rehabilitation. He collaborated in several projects of the RoboCity2030 Consortium. He is currently with the Department of Systems Engineering and Automation, UC3M, where he is involved in teaching activities. His research interest includes assistive and rehabilitation robotics, mechatronics, power electronics, and serious games for health.

PATRICIA SÁNCHEZ-HERRERA BAEZA received the degree in occupational therapy from the Autónoma University of Madrid, Spain, and the Ph.D. degree from the Rey Juan Carlos University of Madrid, Alcorcón, Spain, in 2014, where she is currently a Visitor Professor with the Department of Physical Therapy, Occupational Therapy, Rehabilitation and Physical Medicine. Since 2013, she has been a member of the Motion Analysis, Ergonomics, Biomechanics and Motor Control Laboratory, Faculty of Health Sciences, Rey Juan Carlos University. She has participated in national and international research and development projects. Her research interests include stroke rehabilitation, rehabilitation robots, and cognitive impairment.

ALBERTO JARDÓN HUETE (M'07–SM'13) received the B.Sc. degree in electronics engineering, the master's degree in electrical engineering, and the Ph.D. degree in electric, electronics, and industrial automation engineering from the University Carlos III of Madrid, in 1998, 2002, and 2006, respectively. Since 1997, he has been an active member of the Robotics Lab and has collaborated in the development of the climbing robots ROMA I, ROMA II, and MATS (also named ASIBOT). He was involved with the GEOST-Ciudad Multidimensional, I3CON (EU), and several tunneling and mining projects funded by several industrial clients and European and National Funding. His research is focused on different projects to apply robotics technologies from underground, building, and aerospace industries. He is also responsible for the Assistive Robotics Technologies Lab. He is also focused on the design and development of professional and personal robotic devices for autonomy restoration, such as light-weight service robots, technical aids and the development of applied algorithms, and the design of custom controllers. He holds eight patents. His interests include assistive robotic design, mechatronics, the research in advanced ''user in the loop'' control schemes to improve usability, and the performance of domestic robots. The development of tools to perform this research and the transfer of robotics technology to industry also fit to his priorities.

CARLOS BALAGUER (M'87) received the Ph.D. degree in automation from the Polytechnic University of Madrid (UPM), Spain, in 1983. From 1983 to 1994, he was with the Department of Systems Engineering and Automation, UPM, as an Associate Professor. Since 1996, he has been a Full Professor of the Robotics Lab with the University Carlos III of Madrid. He was the Director of the Department of Systems Engineering and Automation (2006–2007), and the Vice-Rector for Research of the university (2007–2015). His research interests include, but is not limited to, humanoid robotics, robots' design and development, robot control, path and task planning, force-torque control, assistive and service robots, rehabilitation and medical robots, climbing robots, robotics and automation in construction, and human–robot interaction. He has participated in numerous EU projects, since 1989, including the Eureka projects SAMCA, AMR y GEO; Esprit projects ROCCO and CEROS; Brite Project FutureHome; IST project MATS; 6FP IP Projects ManuBuild, I3CON, Tunconstruct; Strep project RobotCWE; 7FP project RoboSpect; and H2020 projects STAMS and BADGER (coordinator). He has published more than 200 papers in journals and conference proceedings, and several books in the field of robotics. He is a member of the IFAC, and the former President of the IAARC (2001–2004). He participates in the European networks EURON and CLAWAR. He is a member of the Editorial Board of *Automation in Construction Journal* (Elsevier). He is the Co-Ordinator of the Madrid Community Universities' Consortium RoboCity2030 on Service Robots (2006–2018). Since 2015, he has been a member of the euRobotics Board of Directors. Since 2016, he has been the Chairman of the Council for Science and Technology of the Community of Madrid. He received several awards, among them for the best book ''*Fundamentos de Robotica*'' edited by McGraw-Hill (1988), the Best Paper of the ISARC'2003 in Eindhoven, The Netherlands, the IMSERSO's Award 2004 for assistive robots research, the Industrial Robot Journal Innovation Award of the Clawar'2005 in London (U.K.), and the Tucker-Hasegawa Award 2006 in Tokyo, Japan, for a major contribution in the field of Robotics and Automation in Construction and FUE's Award 2014 for AIRBUS-UC3M Joint R&D Center. He was the General Chair of the IEEE-RAS Humanoids'2014 conference, and is the General Chair of the IEEE/JRS IROS'2018 to be held in Madrid. He was an Associate Editor of the *IEEE Robotics and Automation Magazine* (2000–2005).