Post-stroke upper limb rehabilitation using virtual reality interventions: Do outcome measures assess extent or type of motor improvement?

Sandeep K Subramanian  
Department of Physical Therapy,  
School of Health Professions  
UT Health San Antonio  
San Antonio, TX, USA  
subramanias3@uthscsa.edu  
0000-0002-5972-1588

Mackenzie K Cross  
Department of Physical Therapy,  
School of Health Professions  
UT Health San Antonio  
San Antonio, TX, USA  
crossmk@livemail.uthscsa.edu

Cole S Hirschhauser  
Department of Physical Therapy,  
School of Health Professions  
UT Health San Antonio  
San Antonio, TX, USA  
hirschhauser@livemail.uthscsa.edu

Vinette BK Johnson  
School of Kinesiology,  
Lakehead University  
Thunder Bay, ON, Canada  
vinbenj@gmail.com

Abstract— Post-stroke upper limb motor improvement continues to remain sub-optimal in a significant proportion of individuals sustaining a stroke. Efforts to enhance UL motor improvement have led to the use of evidence-based interventions including virtual reality technology. The effects of interventions on motor impairments, activity limitations and participation restrictions are commonly assessed using clinical outcomes. Majority of the clinical outcomes focus on the extent of motor improvement (i.e. how much). Information on the type (i.e. how) of recovery can be obtained by using a selected few clinical outcomes and movement pattern kinematic measures. The study objective was to characterize the outcomes used to assess the effects of virtual reality interventions in terms of quantifying the extent and type of upper limb motor improvement. We reviewed the published literature on the effects of virtual reality (VR) based interventions to enhance UL motor improvements. Outcomes from the retrieved studies were initially classified under the appropriate International Classification of Functioning categories. We then categorized the outcomes into those quantifying into type or extent of motor improvement based on existing evidence. We found 100 papers that investigated the effects of virtual reality interventions to enhance post-stroke UL motor improvement. Forty two different outcome measures were used across the 100 studies. Seventeen different outcomes assessed impairments, 16 were used to measure activity limitations and 6 measured participation restrictions and the effects of contextual factors. The Fugl Meyer Assessment, Wolf Motor Function Test and Stroke Impact Scale were most commonly used across the three categories. Of the retrieved 100 studies, 48 used an outcome that considered the type of recovery. Although a smaller proportion, 17 studies included outcomes of movement patterns. The use of outcomes considering the type of recovery is steadily increasing in studies using VR for post-stroke UL rehabilitation.

Keywords — outcomes, kinematics, recovery, compensation, arm

I. INTRODUCTION

Stroke continues to be a global burden and a leading cause of adult disability [1]. Upper limb (UL) hemiparesis is one of the most disabling consequences of a stroke [2]. It often results in an impaired ability to successfully perform daily life activities (ADL) using the more-affected side, which often persists into the chronic stage [3]. Despite the availability of high-quality evidence, achieving maximal levels of UL motor improvement continues to remain a challenge [3]. Efforts to enhance UL motor improvements have led to incorporation of evidence-based interventions including robotics and virtual reality (VR) technologies [4, 5].

Inherent to the successful application of any intervention is the choice of the appropriate outcome measure to assess improvement. In post-stroke UL rehabilitation, outcomes are used to assess the severity of motor impairments, limitations in ADL performance as well as participation restrictions [6]. A wide variety of measures are available to assess issues across all three levels of the International Classification of Functioning (ICF) [7, 8]. On the other hand, the wide variation in the choice of measures also leads to issues in the synthesis of research findings [9]. This has prompted efforts to standardize the use of assessments to be able to better compare the effects of the same intervention across different studies. Recommendations are now available on the most suitable UL measures to be used in studies to enable comparisons [10].

These recommendations [10] suggest using the Fugl-Meyer Assessment (FMA) and Action Research Arm Test (ARAT) as outcomes at the impairment and activity limitation levels of the ICF [11]. Both measures have well established psychometric properties and standardized instructions for test administration.
[12, 13]. However, these measures focus mainly on task–
completion. The movement patterns used to complete the task are not considered. Previous studies have shown that use of compensatory trunk displacement movements explains about 11% (in ARAT) [14] and 52% of the variance in FMA scores [15]. Thus, while the use of these measures quantifies the extent (i.e. how much) of motor improvement, they fail to consider the type (i.e. use of movement patterns).

Use of kinematic measurement can provide information on aspects of movement patterns (joint range of motion, trunk displacement, interjoint co-ordination) and motor performance (movement straightness, speed and precision) [16]. Movement pattern outcomes help in distinguishing between recovery and compensation at the behavioral level [16, 17]. Recovery refers to the reacquisition of pre-morbid pattern of joint movements and rotations used to UL task performance. Compensations can be further subdivided into adaptive and substitutive categories. Adaptive compensations entail use of different joint rotations enabling task completion using the same end effector. Substitutive compensations require the use of a different end effector to achieve the same task.

The use of kinematic motor performance measures as impairment level outcomes is being increasingly encouraged [10, 18]. However, motor performance can be improved using compensatory movement patterns [19, 20]. Thus, consideration of movement pattern outcomes provides a better picture of motor improvement. It can help qualitatively quantify UL motor improvement in response to different post-stroke interventions, including the use of VR.

Two previous studies [21, 22] have reviewed outcomes measures used in studies involving use of VR for post-stroke rehabilitation. These articles included the use of VR for both upper and lower limb rehabilitation and did not focus on movement quality outcomes. Given that post-stroke UL motor improvement continues to be sub-optimal, it is essential to identify the best outcomes to be used with task practice in virtual environments. This will help characterize all aspects of motor improvement (i.e. the extent and the type). In addition, it will be useful in synthesis of research findings using techniques such as meta-analyses. The objective of our study was to characterize the outcomes in studies involving VR interventions in terms of measuring the extent and type of UL motor improvement after a stroke.

II. Methods

A. Literature review

A comprehensive search of studies published in English between 2000-2018 involving adults was conducted by MKC and CSH. Databases searched included Medline, ISI Web of Science and IEEE Xplore. A variety of combinations of MeSH terms and key words including; stroke, cerebrovascular accident, upper limb, rehabilitation, virtual reality, virtual rehabilitation, hemiparesis, arm, outcome, measures and assessments were used. Included articles addressed the effects of multiple sessions of an intervention involving a VR component. We excluded case studies, studies with very small sample sizes (<5 participants) study protocols and systematic reviews. If the preliminary study results were published as part of a conference proceeding and the full paper was subsequently published as a different journal article, we only considered results from the journal article.

B. Data Extraction

We initially extracted information regarding all the outcomes that were assessed in each retrieved article. Then, we grouped the extracted outcomes under appropriate categories of the ICF, based upon available guidelines [23-25]. Some outcomes had items that could be classified under more than one category of the ICF. These outcomes were placed in the category most commonly used in the literature. The outcomes were then classified as those measuring the extent or type of recovery, in reference to existing literature [16, 26].

III. Results

We obtained a total of 100 studies after excluding duplicate citations. The proportions of studies obtained from the different databases using outcomes across the three different ICF categories and contextual factors is provided in Fig. 1. While 93% of the studies employed measures evaluating impairment, limitations in activity performance were assessed in 85 studies. Contextual factors (including measures of quality of life, motivation and mood) were assessed in 13 studies. A total of 10 studies used participation outcomes. The number of studies with the names of the different outcomes used to assess the severity of motor impairment, activity limitations, participation restrictions and contextual factors are presented in Tables 1, 2 and 3 respectively.

A. Motor Impairment outcomes

Twenty different assessments were used across the 86 studies to assess the severity of motor impairment (Table 1). All the studies used ≥1 measure of impairment. Studies using ≥1 measure of impairment evaluated effects of the intervention on function and strength or spasticity as well as kinematic measures. The FMA was the most commonly employed...
outcome, being used in 59 studies. Motor performance kinematic measures were the second most commonly assessed type of outcome, being used in 40 studies. Manual Muscle testing, the Motricity Index, dynamometry as well as pinch and grip strength assessments quantified muscle strength. The joint ranges of motion were assessed using kinematic measurement as well as goniometry. Spasticity was assessed using the Original and Modified Ashworth’s Scale. Additional outcomes used included use of fMRI, Motor Evoked Potentials (MEPs), Brunnstrom’s stages and the Reaching Performance Scale in Stroke (RPSS).

**TABLE 1. OUTCOMES USED TO MEASURE MOTOR IMPAIRMENT**

<table>
<thead>
<tr>
<th>Impairment level assessment</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fugl-Meyer Assessment</td>
<td>59</td>
</tr>
<tr>
<td>Motor performance kinematic measures (movement speed, error, straightness, smoothness)</td>
<td>40</td>
</tr>
<tr>
<td>Strength assessments</td>
<td>30</td>
</tr>
<tr>
<td>Movement pattern kinematic measures</td>
<td>13</td>
</tr>
<tr>
<td>Spasticity</td>
<td>7</td>
</tr>
<tr>
<td>fMRI outcomes</td>
<td>4</td>
</tr>
<tr>
<td>Changes in Motor Evoked Potential after TMS</td>
<td>4</td>
</tr>
<tr>
<td>Motor Function Test</td>
<td>3</td>
</tr>
<tr>
<td>Range of motion using goniometry</td>
<td>3</td>
</tr>
<tr>
<td>Reaching Performance Scale in Stroke</td>
<td>3</td>
</tr>
<tr>
<td>Maximal Forward Reach Distance (cm)</td>
<td>2</td>
</tr>
<tr>
<td>Brunnstrom’s stages</td>
<td>1</td>
</tr>
<tr>
<td>Chedoke McMaster Stroke Assessment</td>
<td>1</td>
</tr>
</tbody>
</table>

**B. Activity Limitation outcomes**

Nine different outcomes were used to measure activity limitations (Table 2). The Wolf Motor Function Test (WMFT) was used most commonly in 35 studies. Hand manipulation and dexterity skills were assessed using 4 different outcomes including the Box and Blocks Test, the Nine Hole Peg Test, the Perdue Pegboard and the Grooved Pegboard Test in 32 studies. Other assessments used less frequently to measure limitations in ADL included the Action Research Arm Test and the Motor Activity Log (15 studies each), Jebsen-Taylor function test (12 studies), Functional Independence Measure (11 studies) the Barthel Index (including the Modified and the Korean Versions – 9 studies), the Chedoke Arm and Hand Activity Inventory (CAHAI – 6 studies). Assessments used in three or less studies included the ABILHAND, Functional tests of the Hemiparetic Upper Extremity, Assessment of Motor and Processing Skills, the Canadian Occupational Performance Measure, Fatigue Severity Scale and the Stroke Upper limb Capacity Scale.

**TABLE 2. OUTCOMES USED TO MEASURE ACTIVITY LIMITATIONS**

<table>
<thead>
<tr>
<th>Activity level assessment</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wolf Motor Function Test</td>
<td>35</td>
</tr>
<tr>
<td>Hand function assessments</td>
<td>32</td>
</tr>
<tr>
<td>Action Research Arm Test</td>
<td>15</td>
</tr>
<tr>
<td>Motor Activity Log</td>
<td>15</td>
</tr>
<tr>
<td>Jebsen-Taylor Test</td>
<td>12</td>
</tr>
<tr>
<td>Functional Independence Measure</td>
<td>11</td>
</tr>
<tr>
<td>Barthel Index</td>
<td>9</td>
</tr>
<tr>
<td>Chedoke Arm and Hand Activity Inventory</td>
<td>6</td>
</tr>
<tr>
<td>ABILHAND</td>
<td>3</td>
</tr>
<tr>
<td>Nottingham Extended Activities of Daily Living</td>
<td>2</td>
</tr>
<tr>
<td>Functional Test of the Hemiparetic Upper Extremity</td>
<td>2</td>
</tr>
<tr>
<td>Assessment of Motor and Processing Skills</td>
<td>1</td>
</tr>
<tr>
<td>Canadian Occupational Performance Measure</td>
<td>1</td>
</tr>
<tr>
<td>Fatigue Severity Scale</td>
<td>1</td>
</tr>
<tr>
<td>Stroke Upper Limb Capacity Scale</td>
<td>1</td>
</tr>
</tbody>
</table>

**C. Participation Restriction and Contextual factor outcomes**

**TABLE 3. OUTCOMES USED TO MEASURE PARTICIPATION RESTRICTION AND CONTEXTUAL FACTORS**

<table>
<thead>
<tr>
<th>Participation level and other assessments</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke Impact Scale</td>
<td>10</td>
</tr>
<tr>
<td>Intrinsic Motivation Inventory</td>
<td>7</td>
</tr>
<tr>
<td>Quality of Life measures</td>
<td>3</td>
</tr>
<tr>
<td>Depression assessments</td>
<td>3</td>
</tr>
</tbody>
</table>

The Stroke Impact Scale was the only outcome used to assess participation restriction in 10 studies. In terms of contextual factors, quality of life, motivation and the presence of depression were assessed (Table 3). The SF-36, Stroke Specific Quality of Life and EuroQol – 5D 3L outcomes helped assess quality of life. The Intrinsic Motivation Inventory was used to assess motivation. Depression was assessed using the
Beck’s Depression Inventory and the Hamilton Depression Rating Scale.

D. Extent vs type of motor improvement

Outcomes that assessed the type of recovery were only used in 48 out of 100 studies. The proportion of studies that used outcomes assessing behavioral recovery and different types of compensations is given in Table 4. Amongst these 48 studies, 29 used outcomes (WMFT and CAHAI) that considered the presence of substitutive compensations while assigning a score. Nine studies used outcomes that assessed behavioral recovery (movement pattern kinematics) as well as accounted for the use of substitutive compensations (WMFT/CAHAI). Two studies used outcomes that evaluated both adaptive (Reaching Performance Scale in Stroke; RPSS) and substitutive (WMFT) compensations. Behavioral recovery using movement kinematics/ranges of motion were evaluated in seven studies. Only one study used movement kinematics and clinical outcomes that assessed the use of both adaptive (RPSS) and substitutive compensations (WMFT).

TABLE 4. OUTCOMES THAT MEASURE THE TYPE OF MOTOR IMPROVEMENT

<table>
<thead>
<tr>
<th>Construct assessed</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substitutive Compensation only</td>
<td>29</td>
</tr>
<tr>
<td>Behavioral Recovery and Substitutive</td>
<td>9</td>
</tr>
<tr>
<td>Behavioral Recovery only</td>
<td>7</td>
</tr>
<tr>
<td>Substitutive and Adaptive Compensation</td>
<td>2</td>
</tr>
<tr>
<td>Behavioral Recovery, Substitutive and</td>
<td>1</td>
</tr>
<tr>
<td>Adaptive Compensation</td>
<td></td>
</tr>
</tbody>
</table>

IV. DISCUSSION

Amongst studies using VR interventions for post-stroke UL motor improvement, we sought to characterize the outcomes in terms of measurement of extent or type of motor improvement. Initially, we initially classified the outcomes based upon the categories of the ICF. Our results indicate that more than 80% of all the studies we retrieved used outcomes assessing participation restrictions and/or effects of contextual factors.

A. ICF Categories

A total of 35 different outcomes were used across all the different ICF categories. All of outcomes have well-established psychometric properties [23-25]. Although quite variable, these numbers are lower than those reported previously [8, 9]. One of the reasons for the low numbers is that while previous reports have focused on the effects of a variety of interventions, our current study focused on the effects of only VR based interventions.

Amongst the 20 different outcomes used to assess motor impairment, the FMA and motor performance kinematic measures were used most commonly. As previously mentioned, international consensus panels [11] recommend the use of these outcomes as measures of motor impairment. Thirty percent of the studies retrieved assessed muscle strength. The use of strength measures needs to be encouraged, especially as they are a quick and feasible assessment. In addition, muscle strength assessment can help make predictions about UL motor improvement capacity [27].

Although an exceedingly small proportion, it is interesting to note that studies are beginning to use outcomes indicative of change in neural pathways in response to VR interventions. Four studies each used fMRI to assess the extent of activation in ipsilesional areas and transcranial magnetic stimulation (TMS) to derive motor map measures and MEPs. Thus, there is a growing interest in the mechanisms of UL motor improvement facilitated using VR interventions. The presence or absence of MEPs on the more-affected side is an additional factor that helps predict the UL motor improvement capacity [27].

The WMFT was the most commonly used activity limitation measure, being used in little more than a third (35%) of all studies. In addition, other studies used UL specific outcomes such as the WMFT, ARAT, MAL and CAHAI. Previous reviews of outcome measures have suggested that these assessments may be most appropriate to assess severity of motor impairment and extent of activity limitations [28, 29]. Additional assessments used include more generic outcomes such as the FIM and Barthel’s Index. The use of more generic outcomes might result in the possibility of the detection of no significant change on these measures.

The use of VR interventions generally involves task-specific practice of UL tasks to enhance motor improvement. For task completion, the nervous system to co-ordinate specific number of joints using the available motor abundance, within the given task-constraints [30]. Previous findings involving task-specific UL exercises have argued that even use of activity level measures such as the MAL may not truly reflect the change in inter-joint and inter-segmental co-ordination [31]. This is attributable to the fact that these outcomes include items involving movements requiring patterns of co-ordination vastly different from what the participants were trained in within virtual environments. Previous study results have demonstrated that task-specific changes are limited to the body-segment that is trained and do not translate to significant change in outcomes such as the Barthel’s Index [5] and the FIM [32]. This can be attributed to the fact that we might not be using the most-suitable measure. Thus, there is a need to judiciously select the outcome to ensure measurement and detection of appropriate change.

Our results also reveal studies are beginning to use outcomes assessing participation restrictions as well as the effect of contextual factors. The inclusion of participation and quality of
life outcomes is a sign of progress, given that exclusion of these outcomes was a routine limitation in prior studies [20]. The inclusion of motivation and depression outcomes should also be encouraged, given that they tend to be confounding factors that can influence the delivery as well as effect of the intervention [31].

B. Type of motor improvement

Forty-eight studies used outcomes assessing the type of motor improvement. Majority of these studies used a measure where individuals scored lower if they used substitutive compensations. Only about a third of these studies (n = 17) used outcomes that assessed behavioral recovery. Behavioral recovery measures are usually obtained using kinematic analyses.

Kinematic assessment generally requires the use of specialized equipment, technical knowledge and may not be feasible in smaller clinics. If available, it is a resource that is highly desirable to use. The recent availability of low-cost devices such as the Kinect have increased the access to such sophisticated measurement. Initial studies reported significant measurement errors [33, 34] with use of the Kinect. However, this issue seems to have been resolved with use of software modifications and filtering parameters [35, 36]. Another viable option to obtain movement pattern kinematic outcomes is using inertial measurement units [37].

VI. REFERENCES


In terms of clinical outcomes, the Reaching Performance Scale in Stroke (RPSS) [38] is one of the few measures that assesses movement patterns. The RPSS is based upon observational kinematics [39], has well established psychometric properties [38, 40] and includes items that assess motor performance and movement patterns. It measures adaptive compensations, as the use of altered movement patterns entails a lower score. The RPSS represents a feasible clinical outcome that quantifies the type of recovery. In the absence of kinematic measurement, the use of the RPSS has been suggested as a suitable alternative outcome [16].

V. CONCLUSION

Majority of the studies that use VR to enhance post-stroke UL rehabilitation use outcomes that assess the extent of motor improvement. However, the proportion of studies using outcomes that assess the type of recovery is steadily increasing. We found that the FMA, WMFT and SIS are the most commonly used outcomes the impairment, activity limitation and participation restriction levels respectively. In addition to the outcomes mentioned above, the inclusion of movement pattern type as kinematic well as motor performance outcomes (if available) or RPSS will provide additional information on the type of recovery. The use of more than one outcome is suggested to capture all facets of motor improvement [26, 41].


