Is the Downs and Black scale a better tool to appraise the quality of the studies using virtual rehabilitation for post-stroke upper limb rehabilitation?

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

Sheena M Caramba
Department of Physical Therapy
School of Health Professions
UT Health San Antonio
San Antonio, USA
caramba@livemail.uthscsa.edu

Oscar L Hernandez
Department of Physical Therapy
School of Health Professions
UT Health San Antonio
San Antonio, USA
hernandezo2@livemail.uthscsa.edu

Quenton T Morgan
Department of Physical Therapy
School of Health Professions
UT Health San Antonio
San Antonio, USA
morganq@uthscsa.edu

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

***Abstract*— We appraised the quality of the Randomized Controlled Trials using virtual reality technology for post-stroke upper limb rehabilitation. We used the Downs and Black scale and PEDRo scales used for study quality appraisal. Correlation analyses revealed that the total scores of the two scales were moderately correlated. When only the items that were similar between two scales were considered, the correlation was high. Preliminary results suggest that the Downs and Black checklist may be a better option to assess quality of studies that use virtual reality technology for post-stroke upper limb rehabilitation.**

Keywords — evidence-based practice, arm, scales, randomized controlled trials,

I. Introduction

Virtual reality (VR) technology is being increasingly used as an intervention to enhance post-stroke upper limb (UL) motor improvement. The availability of a wide variety of low-cost exergaming options has increased the access to VR in clinical settings. Appraising the evidence is an essential part of clinical integration of technology. Critical appraisal of the evidence on the technology provides information that facilitates clinical applicability [1].

Parent randomized controlled trials (RCTs) represent one of the highest levels of evidence [2]. A wide variety of tools are available to appraise the quality of the published RCTs. The most popular amongst these tools is the PEDro scale. Arguably the gold-standard in this field, the PEDro has well established psychometric properties [3]. However, the PEDro scale only includes items that assess the internal validity of the published studies. It does not evaluate power, reporting standards or external validity. These factors are related to treatment implementation and clinical practice [4].

The Downs and Black checklist is a scale which includes items that assess internal and external validity, reporting standards as well as power. Originally developed by Down and Black [5], the scale has been subsequently modified [6], has good psychometric properties [7] and is being increasingly used to appraise quality of randomized and non-randomized studies [8]. The objective of this study was to estimate whether the Downs and Black scale would be a better alternative to appraise the quality of published RCTs using VR for post-stroke UL rehabilitation.

II. Methods

Using standard methodology, MKC and CSH systematically reviewed the literature published in English language published between 2000-2019, with further verification by SMC, OLH and QTM. We included RCTs i)

Note: This is not the official copyright released version of the IEEE proceedings paper. When citing this paper, use the following format: Subramanian S, Caramba S, Hernandez O, Morgan Q, Cross M and Hirschhauser C, "Is the Downs and Black scale a better tool to appraise the quality of the studies using virtual rehabilitation for post-stroke upper limb rehabilitation?", Proc. 13th Int'l Conf. on Virtual Rehab., WG Wright, S Subramanian, G Fluet, M Agmon, RM Proffitt, M Roberts (Eds), Tel Aviv, Israel, 21-24 July 2019.

involving adult participants with stroke with UL hemiparesis and ii) included task-practice in VR environments as part of the experimental group. Study protocols, systematic reviews and studies focusing exclusively on lower limb outcomes or cognitive impairments were excluded. The quality of the retrieved studies was evaluated using both the modified version of the Downs and Black scale scores (total of 28) as well as the PEDro scale (total of 10) by SMC, OLH and QTM. Discrepancies if any, were resolved by SKS.

Scores on the modified Downs and Black checklist were classified as "excellent" (score 24-28), "good" (score 19-23), "fair" (score 14-18), or "poor" (score ≤13) [9]. We ranked total scores obtained on the PEDro scale as good to excellent, fair, or poor (≥6, 4-5, and ≤3, respectively)[10]. Normality of data distribution assumption was verified using the Shapiro-Wilk test. We performed a correlation analysis between the two scores. Strength of the association between the two scores was quantified by the Pearson correlation coefficient. Values of 0.2 - 0.39, 0.4 - 0.79 and ≥0.8 represented mild, moderate and strong levels of correlation respectively [11]. An α value of $p<0.05$ was considered significant.

III. Results

In our preliminary analysis, we retrieved a total of 24 studies. According to the PEDRo scale, we ranked 14 studies as good-to-excellent, 5 studies as fair, and 3 studies as poor. The PEDro score across all studies was 6 ± 1.3 (mean ± SD).

According to the Downs and Black scores, 1 study had an excellent score, 5 studies were good, 15 studies had a fair score and 2 studies were poor. The Downs and Black scale score across all studies was 17 ± 3.7 (mean ± SD).

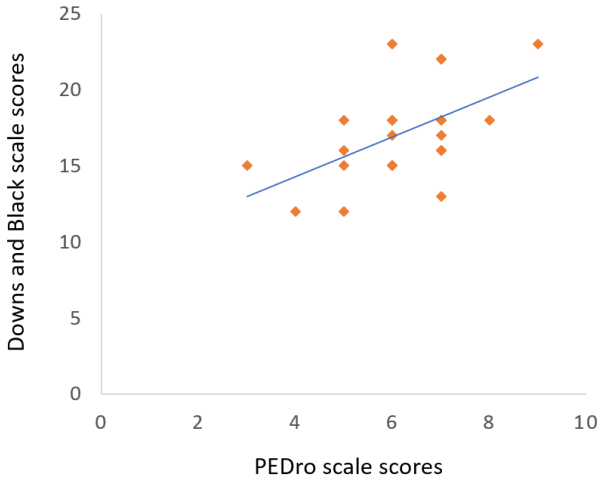

Fig. 1 Results of the correlation analysis between the PEDro and Downs and Black scale scores.

All scores met normality assumptions. Correlation analyses revealed a moderate strength of association between the total scores obtained on both scales ($r = 0.54$, $p < 0.05$; Fig. 1). We also compared the scores from questions which were similar between the two scales. This comparison reveled a high degree of association ($r = 0.7$, $p < 0.05$).

IV. Discussion

Preliminary analyses reveal a moderate strength of correlation between the Downs and Black and PEDro scale total scores. This could be attributed in part to the questions on external validity, reporting standards as well as power in the Downs and Black questionnaire. Our second comparison reveled that the two scores are highly correlated, if we only considered the questions which are similar across the two scales. Thus, using the Downs and Black checklist may be a better option to appraise the quality of the studies using virtual rehabilitation for post-stroke upper limb rehabilitation. Whether the analysis of results across all studies reveals similar or different numbers remains to be estimated.

V. References

[1] D. L. Sackett, "Evidence-based medicine," *Seminars in Perinatology*, vol. 21, pp. 3-5, Feb 1997.

[2] Oxford Center for Evidence Based Medicine. Levels of Evidence [Online]. Available: www.cebm.net, accessed on January 30, 2019.

[3] T. P. Yamato, C. Maher, B. Koes, and A. Moseley, "The PEDro scale had acceptably high convergent validity, construct validity, and interrater reliability in evaluating methodological quality of pharmaceutical trials," *Journal of Clinical Epidemiology*, vol. 86, pp. 176-181, 2017.

[4] S. A. Olivo, L. G. Macedo, I. C. Gadotti, J. Fuentes, T. Stanton, and D. J. Magee, "Scales to assess the quality of randomized controlled trials: a systematic review," *Physical Therapy*, vol. 88, pp. 156-75, Feb 2008.

[5] S. H. Downs and N. Black, "The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions," *Journal of Epidemiology and Community Health*, vol. 52, pp. 377-384, 1998.

[6] S. Morton, C. J. Barton, S. Rice, and D. Morrissey, "Risk factors and successful interventions for cricket-related low back pain: a systematic review," *British Journal of Sports Medicine*, vol. 48, pp. 685-91, Apr 2014.

[7] J. M. Hootman, J. B. Driban, M. R. Sitler, K. P. Harris, and N. M. Cattano, "Reliability and validity of three quality rating instruments for systematic reviews of observational studies," *Research Synthesis Methods*, vol. 2, pp. 110-8, Jun 2011.

[8] S. K. Subramanian and S. S. Prasanna, "Virtual reality and noninvasive brain stimulation in stroke: how effective is their combination for upper limb motor improvement? - A meta-analysis," *PM R*, vol. 10, pp. 1261-1270, Nov 2018.

[9] S. R. O'Connor, M. A. Tully, B. Ryan, J. M. Bradley, G. D. Baxter, and S. M. McDonough, "Failure of a numerical quality assessment scale to identify potential risk of bias in a systematic review: a comparison study," *BMC Research Notes*, vol. 8, p. 224, Jun 6 2015.

[10] Heart and Stroke Foundation Canadian Partnership for Stroke Recovery. Stroke Engine Rating of Evidence [Online]. Available: https://www.strokengine.ca/en/glossary/level-of-evidence/, accessed on January 30, 2019.

[11] H. Coolican, *Research methods and statistics in psychology*, 7th ed. New York, NY: Routledge, 2017.