Reliability and Validity of the Upper Limb Physician’s Rating Scale in Children with Cerebral Palsy

Eun Sook Park, Ji-Woon Joo, Seon Ah Kim, Dong-Wook Rha, and Soo Jin Jung
Department of Rehabilitation Medicine, Severance Hospital, Research Institute of Rehabilitation Medicine, Yonsei University College of Medicine, Seoul, Korea.

Received: October 31, 2013
Revised: March 3, 2014
Accepted: March 7, 2014

Purpose: The Upper Limb Physician’s Rating Scale (ULPRS) is a tool that assesses movement quality of the upper limbs. It is used as an outcome measure after botulinum toxin type A injection in children with cerebral palsy (CP). This study aimed to investigate the reliability and validity of the ULPRS in children with spastic CP.

Materials and Methods: Thirty children with spastic CP (M:F=17:13) aged 5 to 13 years old were recruited. The ULPRS was scored based on recorded videotapes by four physicians on two separate occasions. The Melbourne Assessment of Unilateral Upper Limb Function (MUUL) was scored by an occupational therapist. Intraclass correlation coefficients (ICCs), 95% confidence intervals and weighted kappa statistics were calculated for the scores of ULPRS to obtain interrater and intrarater reliability. The relationship between ULPRS and MUUL was assessed using Pearson correlation coefficients.

Results: The ICCs for the total ULPRS scores were 0.94 between raters and 0.99 to 1.00 within raters. The weighted kappa statistics for subitem scores for the ULPRS ranged from 0.67 to 1.00 within raters and from 0.46 to 0.86 between raters. The relationship between ULPRS and MUUL was strong (Pearson correlation coefficient=0.751; p<0.05).

Conclusion: The results demonstrated the high reliability of the total ULPRS score within and between raters. A significant concurrent validity between ULPRS and MUUL also supports the clinical utility of the ULPRS as an outcome measure of spastic upper limb in children with CP.

Key Words: Upper extremity, botulinum toxin type A injection, outcome measure, Melbourne assessment, unilateral upper limb function

INTRODUCTION

Cerebral palsy (CP) is the most common neurologic disorder that causes chronic motor disability in children.1 The majority of CP cases are of the spastic type, and the upper limb is commonly involved. Spasticity has been considered a main contributor to both impairment of function and the development of deformity. Thus, reduction of spasticity plays an important role in managing children with spastic CP. Botulinum toxin type A (BoNT-A) injection into spastic upper limb muscles has been widely used as an adjunct to conventional therapeutic techniques as a means of reducing muscle spasticity and improving function in the affected arm.2 A reduction of muscle
tone, changes in movement pattern, and also an improvement in upper limb function following BoNT-A injection has been described using various outcome measures.3-7

The Upper Limb Physician’s Rating Scale (ULPRS) is a semi-quantitative modification of the Lower Limb Physician’s Rating Scale. It was designed to assess changes in the movement pattern, focusing on all 3 levels of the arm (palm, forearm, and elbow) to determine whether there is an isolated functional impairment, such as thumb in palm, restricted forearm supination, or a total flexion pattern with thumb in palm, wrist in flexion, forearm supinated, and elbow flexed (Supplementary Table 1, only online).2,8,9

The ULPRS has been recommended by previous studies as an outcome measure to assess functional measures and changes in the movement pattern following BoNT-A injection.2,8,10 The merits of ULPRS are its ease of use in measuring spasticity and speed of assessment by a variety of health care providers without special training. Thus, it may be suitable as an outcome measure for periodic assessments in a busy office setting. However, the reliability and validity of ULPRS have not yet been investigated.

The Melbourne Assessment of Unilateral Upper Limb Function (MUUL) is a reliable and valid tool for measuring the quality of upper limb movement in children aged 5 to 15 years with cerebral palsy.11,12 It is widely used to examine the effectiveness of specific therapeutic interventions. In addition, it was considered the best measure of unimanual capacity in a systematic review article.11 However, to our knowledge, there is no information about the relationship between ULPRS and MUUL. We hypothesized that the ULPRS would be a reliable clinical test for assessing upper limb function in children with CP. Therefore, the aims of this study were: 1) to evaluate the intrarater and interrater reliability of the ULPRS in children with spastic cerebral palsy and 2) to examine the concurrent validity of the ULPRS by investigating its relationship with the MUUL.

MATERIALS AND METHODS

Participants
Thirty children with spastic CP (M:F=17:13) were recruited. The Gross Motor Function Classification System (GMFCS) is widely used to assess the gross motor function in children with CP, and the Manual Ability Classification System (MACS) was developed to evaluate a child’s typical manual performance in daily life.14,15 GMFCS and MACS levels of the participants and the distributions of sex, age, and involvement are shown in Table 1.

Ethical approval was granted by the Institutional Review Board and Ethics Committee of Severance Hospital (#4-2012-0265). Because all of the children in this study were younger than 18 years, informed consent was obtained from the parents of the children for their participation in the study.

Measurements
All of the children were videotaped so that physicians could assess the ULPRS of the upper arm while specific tasks were performed. One physician gave instructions and demonstrated the movements while the other physician recorded the videotape. The scoring of the ULPRS from the videotape was completed by 4 physicians on 2 separate occasions with a time interval of at least one week. The score sheet includes 9 items, and 3-, 4-, and 5-point scales are used to score each component of movement tested.

The MUUL for the affected or more severely affected upper limb was scored by one occupational therapist.

Statistical analysis
Missing values on items scores were handled in the analysis as follows: 1) for single-item summaries of reliability, participants with missing values were omitted, and 2) for analysis based on total scores, mean substitution was used so that for each child, their total score was equal to the average on the non-missing items rescaled to the total number of items recorded.

The intrarater and interrater reliability of the total ULPRS scores were examined using intraclass correlation coefficients (ICCs, based on a two-way random effects model with absolute agreement). To calculate intrarater and interrater reli-

Table 1. Characteristics of Participants

| Characteristic                  | Mean±SD (range) |
|--------------------------------|-----------------|
| Sex (M:F)                      | 17:13           |
| Age (yrs)                      | 8.3±3.0 (5–13)  |
| Involvement (bilateral:unilateral) | 24:6          |
| Dominant hand (right:left)     | 10:20           |
| GMFCS level (n=30, I:II:III:IV:V) | 4:5:8:9:4     |
| In unilateral involvement (n=24, I:II:III:IV:V) | 3:2:1:0:0 |
| In bilateral involvement (n=6, I:II:III:IV:V) | 1:3:7:9:4 |
| MACS level (n=30, I:II:III:IV:V) | 3:14:10:3:0   |
| In unilateral involvement (n=24, I:II:III:IV:V) | 1:4:0:1:0 |
| In bilateral involvement (n=6, I:II:III:IV:V) | 2:10:10:2:0 |

GMFCS, gross motor functional classification system; MACS, manual ability classification system.
The mean ULPRS scores performed by four physicians on two different occasions, the ULPRS score at each reading and the mean score for the MUUL are listed in Table 2. There were no missing values among the data.

The ICCs for the total ULPRS scores were 0.94 between raters and 0.99 to 1.00 within raters, indicating very high intrarater and intrarater reliability (Table 3).

The weighted kappa statistics for ULPRS subitem scores ranged from 0.67 to 1.00 within raters. The k values for intrarater reliability were excellent except for subitems of finger opening, thumb in palm, and associated increase in muscle tone for Rater 3, and active supination in flexion and extension for Rater 4. On the other hand, weighted kappa statistics for subitems of ULPRS ranged from 0.46 to 0.86 between raters. The k values between raters were excellent for only 1 subitem and substantial for 4 ULPRS subitems (Table 4). Fig. 1 displays the differences in the total ULPRS scores between occasions for each rater. The means and 95% limits of agreement (LOA) between occasions were 0.03±1.54 for Rater 1, -0.03±3.37 for Rater 2, -0.13±2.79 for Rater 3, and 0.03±1.54 for Rater 4. As shown in Fig. 1, there were no obvious outliers.

The relationship between the ULPRS and MUUL was strong (Pearson correlation coefficient=0.751; p<0.05).

DISCUSSION

There are various measures to assess upper arm function in...
of administration are crucial for periodic assessments. The ULPRS is a simple tool for assessing the quality of movement and function that can be performed by a variety of health care providers, and its ease of application can be useful in a busy clinical setting. In our study, the ICCs of the total ULPRS score were over 0.9 for both intrarater and interrater reliability, which indicates strong reliability of the total ULPRS score. As far as we know, this is the first study to investigate the reliability of the ULPRS. The ICCs for interrater and intrarater reliability of the total ULPRS score in children with CP, such as the Quality of Upper Extremity Skills Test (QUEST), the MUUL, the Assisting Hand Assessment (AHA) and the Shriners Hospital Upper Extremity Evaluation (SHUEE), which differ in terms of which aspect of upper limb function and which age range they target. Therefore, clinicians may select an upper limb measurement tool to assess a specific aspect of upper limb function. The reliability and validity of an assessment tool is important in effectively measuring the outcomes of a specific therapeutic intervention. In addition, convenience and ease of use are crucial for periodic assessments.

### Table 4. Intrarater Reliability as Assessed by Intra-Class Correlation Coefficients (ICC) with 95% Confidence Intervals (CI) and Intrarater Reliability as Assessed by Weighted Kappa Statistics for Individual Subitem Scores on the ULPRS

| Subitems                              | Intrarater reliability (weighted kappa) | Intrarater reliability (weighted kappa) |
|---------------------------------------|----------------------------------------|----------------------------------------|
|                                       | Rater 1 | Rater 2 | Rater 3 | Rater 4 | Rater 1 | Rater 2 | Rater 3 | Rater 4 |
| Active elbow extension                | 0.79    | 0.90    | 0.95    | 0.87    | 0.89    | 0.77    |
| Active supination in extension         | 0.86    | 0.95    | 1.00    | 0.81    | 0.88    | 0.77    |
| Active supination in flexion           | 0.75    | 0.95    | 0.95    | 0.84    | 0.88    | 0.67    |
| Active wrist dorsiflexion              | 0.71    | 0.95    | 0.90    | 0.94    | 0.90    |
| Wrist dorsiflexion                     | 0.47    | 0.90    | 0.84    | 0.88    | 0.88    |
| Finger opening                         | 0.60    | 1.00    | 0.94    | 0.68    | 1.00    |
| Thumb in palm                          | 0.63    | 1.00    | 0.91    | 0.69    | 0.90    |
| Associated increase in muscle tone     | 0.46    | 0.95    | 0.90    | 0.73    | 1.00    |
| Two-handed function                    | 0.50    | 0.95    | 0.90    | 0.88    | 1.00    |

ULPRS, Upper Limb Physician’s Rating Scale scoring.

![Fig. 1. Repeat score differences for each rater. Mean: repeat score mean; difference: repeat-score difference; dotted line: median of repeat-score difference; middle solid line: mean of repeat-score difference; top and bottom solid line: 95% limits of agreement for the repeat-score difference.](image-url)
our study are compatible with those of QUEST and MUUL in the literature, which are commonly used to assess upper arm function for children with CP.11,22,25

As for the intrarater reliability of the ULP RS item scores, the weighted kappa statistics were overall excellent except for 3 subitems for Rater 3 and 2 subitems for Rater 4. On the other hand, the interrater reliability for ULP RS subitems varied from excellent to moderate agreement. In addition, there were no obvious outliers on the Bland-Altman plot, which was narrow. The LOA indicated that the agreement was strong, with the 95% LOA range within ±3.37 on a 47-point scale. These findings suggest that differences in the ULP RS of more than 3.37 after intervention when performed by the same rater may reflect a true change in function as a result of therapeutic intervention.

The ICCs for each item of the ULP RS are in accordance with the reliability of each score on the MUUL in previous studies, where the ICCs varied from low to high (ranging from 0.37 to 0.91).26,27 Some reports demonstrated that the reliability of the scores on each item of the MUUL were acceptable for all items except items 1 and 2.11,26 In contrast, only 9 of 16 items of the MUUL achieved ICCs over 0.7 in Spirtos, et al.’s report. Although the reason for this significant difference was not clear, differences in the severity of the subjects’ cerebral palsy have been suggested as a possible reason in the Spirtos study. In our study, the weighted kappa statistics for only 4 of 9 items were greater than 0.7. The elbow flexors, pronators, and wrist flexors are commonly targeted for spastic upper limb management with BoNT-A injection,10 and thus, our results indicate that changes in movement patterns after BoNT-A injection into those muscles may be reliably measured with the ULP RS. However, ICCs for the five remaining items on the ULP RS were not adequate for clinical use; therefore, scoring guidelines for clarifying differences between individual scoring criteria are needed.

The MUUL is based on 16 items comprising of tasks that are representative of the most important components of unilateral upper limb function (reach, grasp, release, and manipulation).28 The test is administered by a therapist, with a video recording of the child’s performance obtained for subsequent scoring. The reliability and validity of the MUUL has been well-established in previous studies.12,26,27 According to a recent systematic review article, MUUL is the best scale to measure unimanual functional changes as a result of spasticity management.13 Therefore, the validity of the ULP RS was assessed in relation to the MUUL. The significant positive correlation between total ULP RS and MUUL scores in our study indicates that the two scales may be comparable.

According to the International Classification of Functioning, Disability and Health (ICF), upper limb function can be measured based on body impairment, activity, and participation levels. The MUUL is a tool for assessing the activity level of ICF, whereas the ULP RS has components representing impairment levels of ICF in addition to activity levels. Differences in the way components of ICF were measured may have contributed to the weaker relationship between the ULP RS and MUUL.

Changes in movement pattern and quality after BoNT-A injection are of interest to clinicians involved in the management of spastic upper limbs. From this perspective, the ULP RS may have unique clinical value and thus is often used as an outcome measure following BoNT-A injection in the literature.24,10 The acceptable reliability, reproducibility, and validity of the ULP RS demonstrated in our study support the clinical utility of the ULP RS as an upper arm measure for children with spastic CP following BoNT-A injection into spastic upper limb muscles.

Videotaped evaluation of upper extremity function has several advantages as well as certain disadvantages. One advantage is the ability to observe the same performance several times to capture movement quality more accurately. On the other hand, the procedure is time-consuming, and scoring from video clips using the ULP RS may be difficult depending on the angle of the video. Therefore, intrarater and interrater reliability from videotaped performances of upper extremity function in children with CP may differ from reliability measured live. Comparison of reliability between live and video observation is needed.

In addition, the ULP RS originally aimed to measure post-intervention changes; however, therapeutic interventions were not involved in our study, the effectiveness of the ULP RS in this situation was not examined. Further study is needed to demonstrate the validity of measuring post-intervention changes.

Our study demonstrated high ICC values within and between raters for the total ULP RS score and also a strong relationship between the ULP RS and MUUL. These findings support the clinical utility of ULP RS as an outcome measure after spastic management of the upper limb. However, k values for subitem scores within and between raters varied from excellent to moderate agreements. Further efforts are required to provide scoring guidelines that clarify the differ-
ences between individual scoring criteria for the scores of each item.

REFERENCES

1. Mutch L, Alberman E, Hagberg B, Kodama K, Perat MV. Cerebral palsy epidemiology: where are we now and where are we going? Dev Med Child Neurol 1992;34:547-51.
2. Graham HK, Aoki KR, Autti-Rämö I, Boyd RN, Delgado MR, Gaebler-Spira DJ, et al. Recommendations for the use of botulinum toxin type A in the management of cerebral palsy. Gait Posture 2000;11:67-79.
3. Arner M, Eliasson AC, Nicklasson S, Sommerstein K, Hägglund G. Hand function in cerebral palsy. Report of 367 children in a population-based longitudinal health care program. J Hand Surg Am 2008;33:1337-47.
4. Gschwind C, Tonkin M. Surgery for cerebral palsy: Part I. Classification and operative procedures for pronation deformity. J Hand Surg Br 1992;17:391-5.
5. House JH, Gwathmey FW, Fidler MO. A dynamic approach to the thumb-in palm deformity in cerebral palsy. J Bone Joint Surg Am 1981;63:216-25.
6. Tonkin M, Gschwind C. Surgery for cerebral palsy: Part 2. Flexion deformity of the wrist and fingers. J Hand Surg Br 1992;17:396-400.
7. Zancolli EA, Zancolli ER Jr. Surgical management of the hemiplegic spastic hand in cerebral palsy. Surg Clin North Am 1981;61:395-406.
8. Yang TF, Fu CP, Kao NT, Chan RC, Chen SJ. Effect of botulinum toxin type A on cerebral palsy with upper limb spasticity. Am J Phys Med Rehabil 2003;82:284-9.
9. Koman LA, Mooney JF, Smith BP, Goodman A, Mulvaney T. Management of spasticity in cerebral palsy with botulinum-A toxin: report of a preliminary, randomized, double-blind trial. J Pediatr Orthop 1994;14:299-303.
10. Sätilä H, Kotamäki A, Koivikko M, Autti-Rämö I. Upper limb function after botulinum toxin A treatment in cerebral palsy: two years follow-up of six cases. Pediatr Rehabil 2006;9:247-58.
11. Cusick A, Vasquez M, Knowles L, Wallen M. Effect of rater training on reliability of Melbourne Assessment of Unilateral Upper Limb Function scores. Dev Med Child Neurol 2005;47:39-45.
12. Bourke-Taylor H. Melbourne Assessment of Unilateral Upper Limb Function: construct validity and correlation with the Pediatric Evaluation of Disability Inventory. Dev Med Child Neurol 2003;45:92-6.
13. Gilmore R, Sakzewski L, Boyd R. Upper limb activity measures for 5- to 16-year-old children with congenital hemiplegia: a systematic review. Dev Med Child Neurol 2010;52:14-21.
14. Eliasson AC, Krumlinde-Sundholm L, Röslblad B, Beckung E, Arner M, Ohrvall AM, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. Dev Med Child Neurol 2006;48:549-54.
15. Rosenbaum PL, Palisano RJ, Bartlett DJ, Galuppi BE, Russell DJ. Development of the Gross Motor Function Classification System for cerebral palsy. Dev Med Child Neurol 2008;50:249-53.
16. Laschinger HK. Intraclass correlations as estimates of interrater reliability in nursing research. West J Nurs Res 1992;14:246-51.
17. Mellis C, Peat JK, Williams K, Xuan W. Health science research: a handbook of quantitative methods. Crows Nest, New South Wales: Allen & Unwin; 2001.
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
19. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. Med Care 1992;30:917-25.
20. Morris C, Kunizeczuk JJ, Fitzpatrick R, Rosenbaum PL. Reliability of the manual ability classification system for children with cerebral palsy. Dev Med Child Neurol 2006;48:950-3.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
22. Thorley M, Lannin N, Cusick A, Novak I, Boyd R. Reliability of the quality of upper extremity skills test for children with cerebral palsy aged 2 to 12 years. Phys Occup Ther Pediatr 2012;32:4-21.
23. Krumlinde-Sundholm L, Holmeufur M, Kottorp A, Eliasson AC. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. Dev Med Child Neurol 2007;49:259-64.
24. Davids JR, Sabesan VJ, Ortmann F, Wagner LV, Peace LC, Gideon MA, et al. Surgical management of thumb deformity in children with hemiplegic-type cerebral palsy. J Pediatr Orthop 2009;29:504-10.
25. Sorsdahl AB, Moe-Nilssen R, Strand LI. Observer reliability of the Melbourne assessment of unilateral upper limb function. Dev Med Child Neurol 2008;50:146-51.
26. Randall M, Carlin JB, Chondros P, Reddiough D. Reliability of the Melbourne assessment of unilateral upper limb function. Dev Med Child Neurol 2001;43:761-7.
27. Spirtos M, O’Mahoney P, Malone J. Interrater reliability of the Melbourne Assessment of Unilateral Upper Limb Function for children with hemiplegic cerebral palsy. Am J Occup Ther 2011;65:378-83.
28. Klingels K, De Cock P, Desloovere K, Huenaeerts C, Molenaers G, Van Nuland I, et al. Comparison of the Melbourne Assessment of Unilateral Upper Limb Function and the Quality of Upper Extremity Skills Test in hemiplegic CP. Dev Med Child Neurol 2008;50:904-9.
**Supplementary Table 1. Upper Limb Physician’s Rating Scale**

| Parameters                                                                 | Definitions                              | 0  | 1  | 2  | 3  |
|---------------------------------------------------------------------------|------------------------------------------|----|----|----|----|
| Active elbow extension (normal 180°)                                      | >10° reduction                           | 0  |    |    |    |
|                                                                            | 1 to 10° reduction                      |    | 1  |    |    |
|                                                                            | No reduction                             |    |    | 2  |    |
| Active supination in extension (Elbow extended, forearm supinates)        | None                                     |    |    | 0  |    |
| Mid-position: palm to 90° horizontal                                      | Under mid-position                      |    | 1  |    |    |
|                                                                            | To mid-position                         | 2  |    |    |    |
|                                                                            | Past mid-position                       |    | 3  |    |    |
| Active supination in flexion (elbow flexed at 90° forearm supinates)     | None                                     |    |    | 0  |    |
|                                                                            | Under mid-position                      |    | 1  |    |    |
|                                                                            | To mid-position                         | 2  |    |    |    |
|                                                                            | Past mid-position                       |    | 3  |    |    |
| Active wrist dorsiflexion (forearm supported, active dorsiflexion of wrist) | None                                     |    | 0  |    |    |
| Mid-position: palm level with forearm                                     | Under mid-position                      |    | 1  |    |    |
|                                                                            | To mid-position                         | 2  |    |    |    |
|                                                                            | Past mid-position                       |    | 3  |    |    |
| Wrist dorsiflexion (angle of movement)                                   | Wrist ulnar deviation                   |    |    | 0  |    |
|                                                                            | Wrist radial deviation                  |    | 1  |    |    |
|                                                                            | Neutral                                 | 2  |    |    |    |
| Finger opening                                                            | Only wrist flexion                      |    | 0  |    |    |
|                                                                            | With wrist in neutral position          |    | 1  |    |    |
|                                                                            | With wrist dorsiflexion                 |    | 2  |    |    |
| Thumb in function                                                         | Within palm                             | 0  |    |    |    |
|                                                                            | Pressed laterally against index finger  |    | 1  |    |    |
|                                                                            | Partly assist in grasp                  |    | 2  |    |    |
|                                                                            | Thumb finger grasp possible             |    | 3  |    |    |
|                                                                            | Active abduction                        | 4  |    |    |    |
| Associated increase in muscle tone                                       | In all manipulative functions           |    | 0  |    |    |
|                                                                            | Only with fine motor manipulation       |    | 1  |    |    |
|                                                                            | Only with walking or running            | 2  |    |    |    |
|                                                                            | None                                    | 3  |    |    |    |
| Two handed function                                                       | Poor, no use of hidden function        |    |    |    | 1  |
|                                                                            | Use of all functions, but limited in ADLs |    |    | 2  |    |
|                                                                            | Use of all functions, but not limited in ADLs |    |    | 3  |    |
| Total score                                                               |                                         |    |    |    |    |

ADL, activity of daily living.
Adapted from Graham HK, et al. Gait Posture 2000;11:67-79.  

2