Rating scales for dystonia in cerebral palsy: reliability and validity

E MONBALIU1,2 | E ORTIBUS3 | F ROELENS2,4 | K DESLOOVERE1,5 | J DEKLERCK6 | P PRINZIE2,7 | P DE COCK8 | H FEYS1

1 Department of Rehabilitation Sciences, Katholieke Universiteit Leuven, Belgium. 2 DC GIOI(S), Dominiek Savio Institute, Belgium. 3 Department of Paediatrics, University Hospital Leuven, Belgium. 4 Department of Paediatrics, Heilig Hart Hospital Roeselare, Belgium. 5 Clinical Motion Analysis Laboratory, University Hospital Pellenberg, Belgium. 6 Catholic University College of Bruges-Ostend, Department Health Care, Belgium. 7 Department of Child and Adolescent Studies, Utrecht University, the Netherlands. 8 Faculty of Medicine, Katholieke Universiteit Leuven, Belgium.

Correspondence to Elegast Monbaliu at Dominiek Savio Instituut, Koolskampstraat 24, 8830 Gits, Belgium. E-mail: Elegast.Monbaliu@faber.kuleuven.be

AIM This study investigated the reliability and validity of the Barry–Albright Dystonia Scale (BADS), the Burke–Fahn–Marsden Movement Scale (BFMMS), and the Unified Dystonia Rating Scale (UDRS) in patients with bilateral dystonic cerebral palsy (CP).

METHOD Three raters independently scored videotapes of 10 patients (five males, five females; mean age 13y 3mo, SD 5y 2mo, range 5–22y). One patient each was classified at levels I–IV in the Gross Motor Function Classification System and six patients were classified at level V. Reliability was measured by (1) intraclass correlation coefficient (ICC) for interrater reliability, (2) standard error of measurement (SEM) and smallest detectable difference (SDD), and (3) Cronbach’s alpha for internal consistency. Validity was assessed by Pearson’s correlations among the three scales used and by content analysis.

RESULTS Moderate to good interrater reliability was found for total scores of the three scales (ICC: BADS=0.87; BFMMS=0.86; UDRS=0.79). However, many subitems showed low reliability, in particular for the UDRS. SEM and SDD were respectively 6.36% and 17.72% for the BADS, 9.88% and 27.39% for the BFMMS, and 8.89% and 24.63% for the UDRS. High internal consistency was found. Pearson’s correlations were high. Content validity showed insufficient accordance with the new CP definition and classification.

INTERPRETATION Our results support the internal consistency and concurrent validity of the scales; however, taking into consideration the limitations in reliability, including the large SDD values and the content validity, further research on methods of assessment of dystonia is warranted.

In the past two decades, interest in cerebral palsy (CP) has increased remarkably.1 However, compared with the spastic type of CP, the assessment and treatment of patients with dystonia is still underreported.2 This lack of research is understandable in view of the complexity of dystonia, the difficulty in measuring it,3 and the previous confusing definition and classification of dystonia in CP.4 In 2005, Bax et al.5 introduced a new definition and classification for CP in accordance with that used by the Surveillance of Cerebral Palsy in Europe.4 According to this new definition, dystonia in CP is ‘an abnormal pattern of posture and/or movement which is involuntary, uncontrolled, recurring and occasionally stereotyped. The posture/movement is dominated by hypokinesia characterized by reduced activity with an increased tone tendency’.4,5 Dystonic CP is also called secondary dystonia.1 It is, together with choreoathetosis, part of the dyskinetic CP group with a prevalence of between 6.5%6 and 14.4%7 of the total CP population.

To gain insight into the severity and distribution of dystonia and to guide and delineate therapeutic interventions, reliable and valid assessments are indispensable. Currently, three rating scales are used: the Barry–Albright Dystonia Scale (BADS),8 the Burke–Fahn–Marsden Scale (BFMMS),9 and the Unified Dystonia Rating Scale (UDRS).10 The BFMMS and the UDRS were developed especially for people with primary dystonia, whereas the BADS was modified from the BFMMS in order to measure secondary dystonia, as occurs in CP. Until now, only the BADS has proven to have good interrater reliability in moderate to severe dystonic CP. However, the clinical criteria11 of the scale have not been fully established. Its reliability was not verified in patients with mild dystonia, and the standard error of measurement (SEM) and smallest detectable difference (SDD), as well as internal consistency, have not yet been reported. Further, preliminary results in a recent intervention study in dystonic CP12 suggested that the sensitivity of the BADS is insufficient in moderate to severe CP, affirming similar findings from the work of Butler et al.13 Also, the BADS scale has not yet been correlated with other dystonia scales. However, the BFMS and UDRS have been applied only in patients with primary dystonia,10 and
clinimetric criteria have not been investigated in patients with dystonic CP.

Therefore, the objectives of this study were (1) to investigate the reliability of the BADS, BFMS, and UDRS in children with dystonic CP by determining the intrarater reliability, SEM, SDD, and internal consistency, and (2) to assess the validity by measuring the relationship between the scales (concurrent validity) and by analyzing the content of the three scales (content validity).

METHOD
Participants
Participants older than 4 years with predominant dystonic CP diagnosed by a paediatric neurologist were included. Exclusion criteria were changes in medication for muscle tone within the last 6 months and orthopaedic/neurosurgical interventions within the last year. Ten patients with bilateral dystonic CP participated (see Table I). Their mean age was 13 years 3 months (SD 5y 2mo, range 5–22y). Using the Gross Motor Function Classification System (GMFCS) to determine the severity of CP, one patient each was classified at levels I to IV and six patients were classified at level V. All participants were recruited from special education schools. An attempt was made to recruit a representative sample of children in the different levels of the GMFCS in accordance with the dyskinetic population study of Himmelman et al.14 Ethical approval was obtained from the Committee of Flemish Motor Disability Institutes (Appendix SI, supporting information published online). Informed consent was given by all participants.

Assessment
All participants were assessed using the BADS, BFMS, and UDRS.

The BADS8 evaluates dystonia in eight body regions (Appendix SII, supporting information published online). Each of the scoring criteria for each region are scored from 0 to 4. The maximum total score is 32, calculated by summation of the region scores.

The BFMS9 is subdivided into a movement scale and a disability scale. Only the movement scale was used for this study. The Burke–Fahn–Marsden Movement Scale (BFMMS) evaluates dystonia in nine body regions (Appendix SIII, supporting information published online). The severity of dystonia is evaluated by the provoking factor subscale and is scored on a 4-point ordinal scale. Severity is scored by the severity factor subscale with a score range from 0 to 4. The individual score for each region is the product of the provoking factor and the severity factor. To ‘downweight’ the eyes, mouth, and neck regions, the scores for these areas are each multiplied by 0.5 before summing all region scores to calculate a total score. The maximal total score is 120.

The UDRS10 evaluates dystonia in 14 body areas (Appendix SIV, supporting information published online). The UDRS has a severity and duration rating. The severity factor ranges from 0 to 4. The duration factor is a 9-point ordinal subscale and ranges from 0 to 4 at intervals of 0.5. This factor assesses whether dystonia occurs at rest or in action, and whether it is predominantly at maximal or submaximal intensity. The individual score for each region is the sum of the duration and motor severity factors. The maximal total score of the UDRS is 112, calculated by summing the individual region scores.

Procedure
All participants were videotaped with a standard protocol, both at rest and during activities. The test duration of the protocol varied between 20 and 40 minutes. The formatted protocol was based on the videotaping protocol of the Dystonia Study Group10 (Appendix SV, supporting information published online) and included examination of all regions needed for the three scales. To assess reliability and concurrent validity, two child neurologists (EO, FR) and one physical therapist (EM) independently scored the three scales of the 10 videotaped patients in a randomized order. All raters had several years of clinical experience and were trained in scoring the three scales. They applied the definitions of dystonia as described, according to each of the investigated scales.

Statistical analysis
Interrater reliability for the individual items and total scores was determined by the intraclass correlation coefficient (ICC, Table I: Patient characteristics and neuroimaging findings

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age, y:mo</th>
<th>Sex</th>
<th>GMFCS level</th>
<th>Type</th>
<th>Age, at imaging, y:mo</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19:6</td>
<td>M</td>
<td>III</td>
<td>MRI</td>
<td>16:7</td>
<td>Thalamus, minimal perirolandic lesions</td>
</tr>
<tr>
<td>2</td>
<td>12:6</td>
<td>M</td>
<td>V</td>
<td>MRI</td>
<td>11:9</td>
<td>Basal ganglia, thalamus lesions</td>
</tr>
<tr>
<td>3</td>
<td>13:3</td>
<td>F</td>
<td>I</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>10:0</td>
<td>M</td>
<td>V</td>
<td>MRI</td>
<td>0:9</td>
<td>Basal ganglia, minimal perirolandic lesions</td>
</tr>
<tr>
<td>5</td>
<td>16:3</td>
<td>F</td>
<td>IV</td>
<td>MRI</td>
<td>14:9</td>
<td>Basal ganglia, thalamus lesions</td>
</tr>
<tr>
<td>6</td>
<td>20:8</td>
<td>M</td>
<td>V</td>
<td>CT</td>
<td>0:8</td>
<td>Normal</td>
</tr>
<tr>
<td>7</td>
<td>9:5</td>
<td>M</td>
<td>II</td>
<td>MRI</td>
<td>0:10</td>
<td>Normal</td>
</tr>
<tr>
<td>8</td>
<td>22:0</td>
<td>F</td>
<td>V</td>
<td>CT</td>
<td>1:2</td>
<td>–</td>
</tr>
<tr>
<td>9</td>
<td>10:2</td>
<td>F</td>
<td>V</td>
<td>MRI</td>
<td>6:0</td>
<td>Basal ganglia, thalamus lesions</td>
</tr>
<tr>
<td>10</td>
<td>5:0</td>
<td>F</td>
<td>V</td>
<td>MRI</td>
<td>0:10</td>
<td>Basal ganglia, minimal thalamus lesions</td>
</tr>
</tbody>
</table>

GMFCS, Gross Motor Function Classification System; MRI, magnetic resonance imaging; CT, computed tomography; –, data not available.
2,1) and 95% confidence intervals (CI). ICC values above 0.90 were considered as excellent, between 0.75 and 0.90 as good, and less than 0.75 as poor to moderate.\textsuperscript{15} We considered ICC values between 0.60 and 0.75 as moderate and less than 0.60 as poor. The SEM and the SDD were calculated by using the following formulae:\textsuperscript{16} \(\text{SEM} = \text{SD} \times \sqrt{1 - \text{ICC}}\) and \(\text{SDD} = \text{SEM} \times 1.96 \times \sqrt{2}\).

Internal consistency was determined by Cronbach’s alpha (\(\alpha\)).\textsuperscript{15} The correlation between the total scores of the BFMM, UDRS, and BADS was assessed by the Pearson correlation coefficient.\textsuperscript{17} Statistical analysis was performed with SPSS 16.0 (SPSS Inc., Chicago IL, USA).

**RESULTS**

**Interrater reliability**

The BADS total score showed good interrater reliability, with an ICC of 0.87 and CI between 0.67 and 0.96 (Table II). The ICC values for the eight items ranged from 0.39 to 0.76, with moderate to good reliability for the limbs and trunk and poor reliability for the regions of eyes, mouth, and neck.

Results for the interrater reliability of the BFMM are presented in Table III. The total score of the provoking factor had a moderate reliability, with an ICC value of 0.64 (95% CI 0.28–0.88). In contrast, good interrater reliability was found for the total score of the severity factor and the total score of the multiplication of the subscales, with ICC values of 0.89 (CI 0.72–0.97) and 0.86 (CI 0.66–0.96) respectively. The ICC coefficients for the items of the provoking factor were poor for seven out of nine items, with the lowest reliability for the limbs. In general, the severity subscale items showed better reliability. Only two items showed poor reliability: the eyes and trunk. Reliability of the other items was moderate to good. Compared with the severity factor, the reliability for the multiplication of the subscales was moderate to good for eight of the nine items.

Similar reliability coefficients were found for the UDRS total scores (Table IV), with an ICC of 0.74 (95% CI 0.43–0.92) for the duration factor and an ICC of 0.79 for both the motor severity factor and the summation of both subscales (95% CI 0.51–0.94 and 0.52–0.94 respectively). For the duration factor items scores, one body region showed a good reliability (neck), two regions showed moderate reliability (lower face and jaw/tongue), and all other body regions had poor reliability. For the motor severity factor, nine items had poor ICC values, three items had moderate reliability (lower face, neck, and left distal leg), and two items good reliability (jaw/tongue and larynx). The items of the summation of the subscales showed 10 ICC values with poor reliability, two with moderate reliability (jaw/tongue and larynx), and two with good reliability (lower face and neck).

**Standard error of measurement and smallest detectable difference**

A high SEM was found: 6.36% for the total BADS score, 9.88% for the BFMM total score, and 8.89% for the total UDRS score. The SDD was also high: 17.72% for the BADS and 27.39% and 24.63% for the BFMM and the UDRS respectively.

**Internal consistency**

Among the three raters, a high level of internal consistency was found for the BADS, with Cronbach’s alpha coefficients ranging from 0.87 to 0.91 (Table V). Similarly, the consistency coefficient of the three raters ranged from 0.83 to 0.94 for the...
Concurrent validity
Pearson’s correlation revealed high associations between the total scores of the BADS and the BFMMS: 0.93 (95% CI 0.73–0.98) for rater 1, 0.95 (95% CI 0.79–0.99) for rater 2, and 0.95 (95% CI 0.69–0.98) for rater 3. Coefficients between the BADS and UDRS were 0.89 (95% CI 0.60–0.97), 0.84 (95% CI 0.6–0.96), and 0.86 (95% CI 0.51–0.97) for rater 1, 2, and 3, respectively. Correlation of the BFMMS and UDRS showed coefficients of 0.87 (95% CI 0.53–0.97) for rater 1, 0.88 (95% CI 0.56–0.97) for rater 2, and 0.93 (95% CI 0.73–0.98) for rater 3; p values were lower than 0.001, except for the correlation between BADS and UDRS for rater 2 (p=0.002).

Content validity
In clinical practice, the presence of dystonia in different body regions can be assessed (1) at rest and during activity, and in terms of (2) duration, (3) amplitude, and (4) influence on functional activities.

The BADS9 and BFMMS9 describe dystonia in eight and nine body regions but do not differentiate between proximal and distal parts of the limbs. The UDRS10 includes ratings for 14 body regions and describes the proximal and distal parts of the arms and legs separately.

The BADS does not differentiate between rest and activity, and it includes within the item descriptions duration, amplitude, and influences on functional activities in a variable and inconsistent way.

The provoking factor of the BFMMS assesses the presence of dystonia at rest and during activity, but combined in one score in a hierarchical way such that the presence of dystonia at rest has the greatest influence on the score. The severity factor includes the evaluation of duration and/or amplitude but in an inconsistent way, depending on the body region. Despite the fact that the influence of dystonia on functional activities is judged in the BFM disability scale, this aspect is also scored in the severity factor.

In the UDRS, the duration factor does not differentiate between rest and activity, and it expresses dystonia as a percentage of duration and amplitude, combined within one item. The score content of the motor severity factor usually includes a measure of amplitude described as a percentage and sometimes a velocity description of the dystonic movement.


total scale and subscales of the BFMMS and from 0.90 to 0.95 for the UDRS.

Table IV: Intraclass correlation coefficients (ICC) with 95% confidence intervals (CI) between raters for the Unified Dystonia Rating Scale

<table>
<thead>
<tr>
<th>Body region</th>
<th>Duration factor ICC</th>
<th>95% CI</th>
<th>Motor severity factor ICC</th>
<th>95% CI</th>
<th>Duration factor × motor severity factor ICC</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eyes</td>
<td>0.42</td>
<td>0.02 to 0.78</td>
<td>0.44</td>
<td>0.04 to 0.79</td>
<td>0.51</td>
<td>0.11 to 0.82</td>
</tr>
<tr>
<td>Lower face</td>
<td>0.71</td>
<td>0.38 to 0.91</td>
<td>0.63</td>
<td>0.26 to 0.88</td>
<td>0.78</td>
<td>0.50 to 0.93</td>
</tr>
<tr>
<td>Jaw/tongue</td>
<td>0.64</td>
<td>0.28 to 0.88</td>
<td>0.81</td>
<td>0.56 to 0.94</td>
<td>0.62</td>
<td>0.25 to 0.88</td>
</tr>
<tr>
<td>Larynx</td>
<td>0.17</td>
<td>−0.17 to 0.63</td>
<td>0.83</td>
<td>0.83 to 0.59</td>
<td>0.63</td>
<td>0.26 to 0.88</td>
</tr>
<tr>
<td>Neck</td>
<td>0.76</td>
<td>0.47 to 0.93</td>
<td>0.74</td>
<td>0.43 to 0.92</td>
<td>0.81</td>
<td>0.57 to 0.94</td>
</tr>
<tr>
<td>Trunk</td>
<td>0.41</td>
<td>0.01 to 0.78</td>
<td>0.54</td>
<td>0.15 to 0.84</td>
<td>0.49</td>
<td>0.09 to 0.81</td>
</tr>
<tr>
<td>Arm, right proximal</td>
<td>0.38</td>
<td>0.02 to 0.76</td>
<td>0.41</td>
<td>0.02 to 0.78</td>
<td>0.51</td>
<td>0.12 to 0.83</td>
</tr>
<tr>
<td>Arm, right distal</td>
<td>0.46</td>
<td>0.06 to 0.80</td>
<td>0.58</td>
<td>0.20 to 0.86</td>
<td>0.59</td>
<td>0.22 to 0.86</td>
</tr>
<tr>
<td>Arm, left proximal</td>
<td>0.32</td>
<td>−0.06 to 0.73</td>
<td>0.5</td>
<td>0.11 to 0.82</td>
<td>0.57</td>
<td>0.19 to 0.85</td>
</tr>
<tr>
<td>Arm, left distal</td>
<td>0.3</td>
<td>−0.08 to 0.72</td>
<td>0.49</td>
<td>0.09 to 0.81</td>
<td>0.47</td>
<td>0.07 to 0.81</td>
</tr>
<tr>
<td>Leg, right proximal</td>
<td>0.44</td>
<td>0.05 to 0.79</td>
<td>0.35</td>
<td>−0.04 to 0.74</td>
<td>0.42</td>
<td>0.03 to 0.78</td>
</tr>
<tr>
<td>Leg, right distal</td>
<td>0.31</td>
<td>−0.08 to 0.72</td>
<td>0.48</td>
<td>0.09 to 0.81</td>
<td>0.39</td>
<td>0.00 to 0.77</td>
</tr>
<tr>
<td>Leg, left proximal</td>
<td>0.41</td>
<td>0.02 to 0.78</td>
<td>0.26</td>
<td>−0.11 to 0.69</td>
<td>0.37</td>
<td>−0.02 to 0.75</td>
</tr>
<tr>
<td>Leg, left distal</td>
<td>0.29</td>
<td>−0.09 to 0.71</td>
<td>0.6</td>
<td>0.22 to 0.87</td>
<td>0.44</td>
<td>0.05 to 0.79</td>
</tr>
<tr>
<td>Total</td>
<td>0.74</td>
<td>0.43 to 0.92</td>
<td>0.79</td>
<td>0.51 to 0.94</td>
<td>0.79</td>
<td>0.52 to 0.94</td>
</tr>
</tbody>
</table>

Table V: Internal consistency (Cronbach’s alpha) of the Barry–Albright Dystonia Scale (BADS), the Burke–Fahn–Marsden Movement Scale (BFMMS), and the Unified Dystonia Rating Scale (UDRS)

<table>
<thead>
<tr>
<th>Rater</th>
<th>BADS</th>
<th>BFMMS</th>
<th>UDRS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PF</td>
<td>SF</td>
<td>PF × SF</td>
</tr>
<tr>
<td>Rater 1</td>
<td>0.87</td>
<td>0.83</td>
<td>0.93</td>
</tr>
<tr>
<td>Rater 2</td>
<td>0.91</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>Rater 3</td>
<td>0.9</td>
<td>0.93</td>
<td>0.91</td>
</tr>
</tbody>
</table>

PF, provoking factor; SF, severity factor; DF, duration factor; MSF, motor severity factor.
descriptions used in the BADS, BFMMS, and UDRS combine the body function and activity level of the International Classification of Functioning, Disability and Health (ICF). 18

**DISCUSSION**

The first aim of this study was to assess the reliability of the BADS, BFMMS, and UDRS.

Interrater reliability for the total BADS score was high. The item scores showed moderate reliability for the limbs and trunk, whereas the reliability for the eyes, mouth, and neck was poor. The low ICC values might be caused by too many dystonia characteristics in the item score description. Similarly, Barry et al. 8 found good reliability for the total score and low reliability for the eyes, mouth, and neck. Our results for the limbs and trunk were generally higher than the results of Barry et al. This could be explained by our higher spread of participants in the GMFCS levels, resulting in a higher variability. 15

The total score of the BFMMS as well as the severity factor showed good reliability. Also, moderate to high ICC coefficients were found for almost all items of the severity factor and for the BFMMS subscale multiplication score. In contrast, most of the items of the provoking factor displayed poor reliability, hence resulting in a clearly lower ICC value for the provoking factor total score. This discrepancy could be explained by the hierarchical construct of the provoking factor scoring dystonia presence at rest and during activity. Inspection of the data revealed that, when dystonia was present, nearly all patients showed dystonia at rest. This resulted in maximum scores of 3 and 4, leading to a lower variability and lower ICC values. 15

The ICCs for the UDRS total score and the two subscales were moderate to good, but lower than the ICCs for the BADS and BFMMS. The subscale items demonstrated, in general, poor reliability. The reliability only of the body regions of lower face, jaw/tongue, and neck was sufficient on both subscales. The very poor reliability of the duration factor items may be caused by the 9-point ordinal scale and the description of two dystonia characteristics within one score. This detailed score range seems too difficult to judge reliably. In-depth analysis of the results revealed that the poor reliability of the motor severity factor items may be because of the difficulty in judging the exact percentage of amplitude.

The interrater reliability of the BFMMS and UDRS has not previously been assessed in patients with primary dystonia. 10 Comella et al. 10 found lower ICC values for the total scores, but good reliability for the BFMMS total score (ICC 0.78) and moderate reliability for the UDRS (ICC 0.71). Item scores were not reported.

The SEM and SDD for the BADS were, respectively, 6.36% and 17.72%. In clinical use, this means that a score difference of 17.72% is necessary to be sure that a ‘true’ improvement has occurred, rather than the difference being due to measurement errors. 19 In the case of the BFMMS and the UDRS, the values of SEM and SDD were even higher. Until now, the SEM or SDD of these scales has not been reported. In other studies, SDD values for measurement scales of upper limb function in children with CP have varied between 9% and 13%. 20 Our results were generally higher. This suggests that the sensitivity of scales to detect true change in longitudinal follow-up or after intervention is lower than previously reported.

The internal consistency of all total and subscale scores on all three scales was high for all three raters. These results indicate a robust rating construct in measuring secondary dystonia. 15 Similar results were found by Comella et al. 10 for the BFMMS and UDRS in patients with primary dystonia. The internal consistency of the BADS has not previously been calculated.

Our second aim was to assess the validity of the scales. As expected, the total scores of the three dystonia scales showed a close relationship because the BADS and UDRS were modified from the BFMMS. Similar results were found by Comella et al. 10 for the BFMMS and UDRS in patients with primary dystonia. Our results for the BADS would therefore support the concurrent validity of the scales in patients with CP.

Content analysis revealed that the three scales included several dystonia characteristics over the most important body regions. However, the items are a combination of several different dystonia characteristics within one score (e.g. duration, amplitude). This hampers the study of the characteristics of dystonia in detail and the ability to gain a deeper insight into the phenomenon of dystonia in patients with CP. In addition, the use of combined characteristics may limit the sensitivity of the scale. Visual inspection of the raw data indeed revealed that the highest scores were awarded in many cases, even for patients classified at GMFCS levels I to III.

Further content analysis revealed that many score descriptions do not accord sufficiently with the definition of dystonia in CP recently described by the Surveillance of Cerebral Palsy in Europe, which clearly differentiates between dystonia and choreoathetosis in CP; dystonia is dominated by reduced activity with an increased tone tendency, whereas choreoathetosis is defined as an abnormal pattern of posture and/or movement which is involuntary, uncontrolled, recurring, and occasionally stereotyped, and is dominated by increased activity with a decreased tone tendency. 4 This clear distinction is not incorporated in the investigated scales. As all three raters reported that dystonia and choreoathetosis were often simultaneously present in the patients, it may have been difficult to distinguish dystonia and choreoathetosis, which, in turn, may have compromised the reliability of several items.

Finally, the movement descriptions used in the BADS, BFMMS, and UDRS combine the body function and activity level of the ICF model (e.g. the duration of dystonia and its impact on mobility are combined in several score descriptions), whereas the ICF model recommends clearly separated assessments. 18 The items often measure the influence of dystonia on functional activities. However, other factors can influence functional activity. Thus, it is relevant to examine the function and activity level in another entity.

This study warrants some critical reflections. Despite the high ICC values (and Pearson’s correlation) obtained, it should
be noted that the small sample size resulted in relatively wide CI. Larger studies are needed to obtain more precise estimates. The sample had a negative skewed distribution based on the GMFCS, which is inherent to this population. However, this distribution skewness may be a contributor to lower reliability. Despite this, an attempt was made to include participants with different levels of motor function according to the GMFCS. In addition, some of the patients showed spasticity features, which may be a confounding factor. This should be elaborated in further studies, which should also evaluate test–retest reliability.

This was the first study to assess several clinimetric criteria of the three current dystonia rating scales in patients with CP. The results yield important information but also show that further study is required on the assessment methods of dystonia in CP. Our study would, therefore, suggest the need for a new scale addressing both dystonia and choreoathetosis in patients with dyskinetic CP. A new scale that (1) differentiates choreoathetosis and dystonia, and (2) sufficiently discriminates dystonia and spasticity will allow further study into the natural history of dystonia and choreoathetosis over time and ascertain whether this can be influenced by therapeutic interventions.

CONCLUSION

In this study, the interrater reliability of the BADS, BFMM, and UDRS total scores was found to be moderate to good and internal consistency in measuring dystonia in patients with CP was high. The high level of association between the three scales supports the concurrent validity. On the other hand, several items of the BADS and BFMM, and most items of the UDRS, showed low interrater reliability. The high SEM and SDD reduce the sensitivity of the scales for clinical use. Further limitations are the insufficient accordance with the new CP definition and classification and the amalgamation of several levels of the ICF model. Thus, further study into the assessment of dystonia in patients with CP is warranted.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Marguerite-Marie Delacroix Foundation. We thank all participants. Special thanks are due to Marc Vermandere, Joke Vercruyssse, Katrien Brans, Inge Aelbrecht, and the Committee of Flemish Motor Disability.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SI: Committee of Flemish Motor Disability Institutes.

Appendix SII: The Burke–Fahn–Marsden (BFM) scale: movement scale.

Appendix SIII: The Burke–Fahn–Marsden (BFM) scale: movement scale.

Appendix SIV: The Unified Dystonia Rating Scale (UDRS) revised.

Appendix SV: Dystonia Study Group videotape examination protocol.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author of the article.

REFERENCES