Reviewer’s report

Title: Observational Skills Assessment Score: reliability in measuring amount and quality of use of the affected hand in unilateral Cerebral Palsy

Version: 2 Date: 27 May 2013

Reviewer: Ann-Kristin Gunnes Elvrum

Reviewer’s report:

This study aimed to measure intra- and inter-rater reliability, and test-retest reliability of the newly developed Observational Skills Assessment Score (OSAS) in children with unilateral spastic cerebral palsy (CP). A convenient sample of 32 children with unilateral spastic CP was included with 16 children in each of the following age groups: 2.5-6 (younger age group) and 12-16 (older age group) years of age. Three age specific bimanual tasks were performed in both groups. All children performed one measurement session. Two measurements sessions were performed six weeks apart for all children in the older age group and two weeks apart for ten children in the younger age group. A total of three experienced therapists participated in testing and eight physiotherapists or occupational therapists performed the ratings. The authors found that OSAS appears to be a reliable assessment tool of quality of use of the affected hand, with good agreement between repeated measures.

There are some shortcomings in this manuscript that need to be addressed.

Major compulsory revisions:
1. More precise descriptions with references are needed in the Background to clarify how the authors define capacity and performance. This will have implications for the rest of the manuscript.
2. The rational for and aim of OSAS is unclear. It says in the Background, 3rd paragraph that OSAS combines both aspects of unilateral capacity (MUUL) and bimanual performance (AHA), but it is unclear what is meant by this. Is both capacity and performance assessed, or do the authors refer to the terms unilateral and bilateral? The description of OSAS in the Method adds to the confusion. Is the child “forced to use the hand” (stated in the 1st paragraph, Method section) or is “spontaneous use” assessed (stated in the 3rd paragraph, Methods section)? A more thoroughly and detailed description of the scope, tasks and administration of the tasks of OSAS would be helpful.
3. It is confusing to read the description of various numbers of participants and raters for different reliability measures in the Method. Is it possible to make a flow chart to clarify this?
4. The Guidelines for Reporting Reliability and Agreement Studies (GRRAS) are referred to in this manuscript, which is good. However, when comparing this manuscript with the GRRAS guidelines there are some shortcomings that need
to be addressed. For example the sampling method for both rated subjects and raters are not clearly stated. The sampling method used in this study allows for blinding, but it is not clear whether blinding was actually used. Furthermore it is unclear what kind of training was provided to the raters and if the rating was done independently.

5. In this study both quality of use of the affected hand (categorical data) and amount of use (continuous data?) were scored. Has it been considered to use different statistical methods for the various data?

6. There is a lot of information in the Results and a lot of numbers and information in the figures and tables. Maybe it could be considered to separate the reporting of the reliability of OSAS for the two age groups since different tasks are performed?

7. Limitations of the study are not clearly stated in the Discussion.

Minor Essential Revisions:

1. 1st paragraph, Background: A more precise description of hand function problems in children with unilateral CP could be to include the motor component of the impairment with additional references.

2. 3rd paragraph, Background: Which tasks on the MUUL is not suitable for the assisting hand (except the drawing task)? Is this the authors’ impression or do you have a reference for this statement?

3. 3rd paragraph, Background: What is meant by “The task was chosen to match the children’s level of spatial insight and praxis to prevent the influence of cognitive aspects” and how was this determined? This is also mentioned in the Discussion (6th paragraph) but no references are given.

Discretionary Revisions:

1. BoNT-A is a more commonly used abbreviation for Botulinum toxin A than btA.

2. 2nd paragraph, Background: Maybe it is not necessary to list all the instruments included in the review by Gilmore et al (2010), but merely refer to the conclusion?

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests.