

Clinical pain research

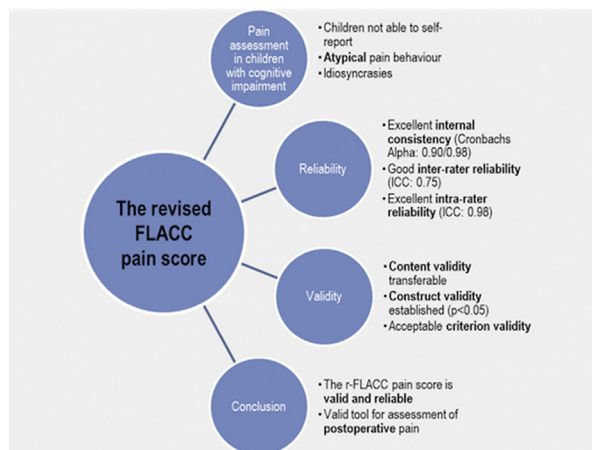
The revised FLACC score: Reliability and validation for pain assessment in children with cerebral palsy

Line Kjeldgaard Pedersen ^{a,b,*}, Ole Rahbek ^{a,b}, Lone Nikolajsen ^c, Bjarne Møller-Madsen ^{a,b}^a Department of Children's Orthopaedics, Aarhus University Hospital, Nørrebrogade 44, 8000 Aarhus C, Denmark^b Orthopaedic Research Laboratory, Department of Orthopaedic Surgery Aarhus University Hospital, Nørrebrogade 44, 8000 Aarhus C, Denmark^c Danish Pain Research Center, Aarhus University Hospital, Nørrebrogade 44, 8000 Aarhus C, Denmark

HIGHLIGHTS

- Management of pain in children with CP not able to self-report is difficult.
- Postoperative pain with atypical pain behaviours and idiosyncrasies is common.
- We evaluated validity and reliability of the translated r-FLACC pain score.
- Excellent internal consistency and intra-rater reliability were found.
- Content, construct and criterion validity were established.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 9 May 2015

Received in revised form 24 June 2015

Accepted 26 June 2015

Available online 24 July 2015

Keywords:

Pain assessment
Cerebral palsy
r-FLACC
Validation
Reliability

ABSTRACT

Background and aims: Pain in children with cerebral palsy (CP) is difficult to assess and is therefore not sufficiently recognized and treated. Children with severe cognitive impairments have an increased risk of neglected postoperative, procedural and chronic pain resulting in decreased quality of life. The r-FLACC (revised Face, Legs, Activity, Cry and Consolability) pain score is an internationally acclaimed tool for assessing pain in children with CP because of its ease to use and its use of core pain behaviours. In addition the r-FLACC pain score may be superior to other pain assessment tools since it includes an open-ended descriptor for incorporation of individual pain behaviours. The COSMIN group has set up three quality domains, which describe the quality of Health-Related Patient-Reported Outcomes (HR-PROs). These are reliability (internal consistency, reliability and measurement error), validity (content validity, construct validity and criterion validity) and responsiveness. The r-FLACC score has only been assessed for reliability and validity in the original English version by the developers of the score. The aim of this study is to assess reliability and validity of the r-FLACC pain score for use in Danish children with CP.

Methods: Twenty-seven children aged 3–15 years old with CP were included after orthopaedic surgery. Two methods for assessment of postoperative pain were used. Pain intensity was assessed by r-FLACC, with a 2 min standardized video recording of the child, and the Observational Visual Analogue Score (VAS-OBS) assessed by the parents. The COSMIN checklist was used as a guideline in the reliability and validity testing of the r-FLACC score.

DOIs of original articles: <http://dx.doi.org/10.1016/j.sjpain.2015.08.004>, <http://dx.doi.org/10.1016/j.sjpain.2015.06.005>

* Corresponding author at: Department of Children's Orthopaedics, Aarhus University Hospital, Nørrebrogade 44, 8000 Aarhus C, Denmark. Tel.: +45 20718424; fax: +45 89494150.

E-mail addresses: line.kjeldgaard.pedersen@clin.au.dk, linekjpe@hotmail.com (L.K. Pedersen), ole0rahbek@gmail.com (O. Rahbek), lone.nikolajsen@clin.au.dk (L. Nikolajsen), bj.moma@gmail.com (B. Møller-Madsen).

<http://dx.doi.org/10.1016/j.sjpain.2015.06.007>

1877-8860/© 2015 Scandinavian Association for the Study of Pain. Published by Elsevier B.V. All rights reserved.

Results: Reliability was supported by three measurement properties. Internal consistency was excellent with a Cronbachs alpha of 0.9023 and 0.9758 (two raters). A factor analysis of the subgroups in the r-FLACC score showed unidimensionality. A test–retest showed excellent intra-rater reliability with an intraclass correlation (ICC) of 0.97530. Inter-rater reliability was acceptable with an ICC of 0.74576. Validity was supported by three measurement properties. Content validity was tested by the originators of the r-FLACC. Construct validity was supported by a significant increase in r-FLACC scores following surgery ($n=17$; difference 2.23; $p=0.0397$). Criterion validity was acceptable with Pearson's correlation coefficients of 0.76 and 0.59 when comparing r-FLACC scores and VAS-OBS scores.

Conclusion and implications: This study benefits from a systematical approach to the validation and reliability parameters by using the COSMIN checklist as a guideline. It is evident that the r-FLACC pain score maintains its psychometric properties after translation. In conclusion, the r-FLACC pain score is valid and reliable in assessing postoperative pain in children with CP not able to self-report pain. With the r-FLACC pain score clinicians have a valid tool for assessing postoperative pain, hence increasing the quality of pain management in children with CP. In addition the validated r-FLACC score has the potential for use in interventional research regarding pain management in this vulnerable group of patients. Future perspectives include validation of the r-FLACC score for procedural and chronic everyday pain and implementation into daily practice.

© 2015 Scandinavian Association for the Study of Pain. Published by Elsevier B.V. All rights reserved.

1. Introduction

Management of pain in children with cerebral palsy (CP) is problematic and due to the complexity of assessment often not adequately treated. Postoperative pain in children with CP is assessed more infrequently than in cognitively healthy children for which reason less analgesia is administered [1]. Management of postoperative pain, especially in the most affected children, is a concern, since a positive proportional relationship exists between the severity of CP, level of cognitive impairment and number of surgical procedures needed. Recent studies report, that a high proportion of children with CP have frequent episodes of pain resulting in decreased quality of life [2,3]. Until present study no validated pain assessment tools for children with CP have been available in Denmark.

The r-FLACC (revised Face, Legs, Activity, Cry and Consolability) pain score is an internationally acclaimed tool for assessing pain in children with CP unable to self-report because of its ease to use, its use of core pain behaviours and its clinical utility. In addition the r-FLACC pain score is superior to other pain assessment tools since the revision introduced an open-ended descriptor for incorporation of typical and atypical individual pain behaviours [4–6]. Children with cognitive impairment may have atypical pain behaviour due to idiosyncrasies masking the typical expression of pain among others laughing, singing, clapping of hands, anger, aggressiveness and self-injury [4,7,8]. The original FLACC score showed low subgroup reliability necessitating the revision. Subsequent reliability and validity testing of the revised score show excellent agreement in all of the 5 categories [4,9].

As with most of the validated pain assessment tools, the r-FLACC is developed in English and has not been validated following translation. Other pain assessment tools for children with CP validated for postoperative pain assessment include the Individualized Numeric Rating Scale (INRS) [10] with established construct validity and inter-rater reliability and the Noncommunicating Children's Pain Checklist-Postoperative Version (NCCPC-PV) [11,12] with established construct validity, inter-rater reliability and internal consistency. Good psychometric properties were found for the Paediatric Pain Profile (PPP) [13] for assessment of everyday pain with established content and construct validity, inter-rater reliability and internal consistency and for the Echelle Douleur Enfant San Salvador (DESS) [14] for procedural pain assessment with established construct validity [15].

The quality of Health-Related Patient-Reported Outcomes (HR-PROs) is best described by three quality domains by the COSMIN group [16]. *Reliability* (internal consistency, reliability and

measurement error) estimates the extent to which scores for patients who have not changed are the same for repeated measurements under several conditions. *Validity* (content validity, construct validity and criterion validity) refers to the instrument actually measuring the construct it is designed to measure. *Responsiveness* refers to the probability of correctly identifying an adequate and valid reflection of changes between measurements [17–20].

To the best of our knowledge the reliability and validity of the r-FLACC pain score have never been published by other than the developers. The aim of this study is to assess reliability and validity of the r-FLACC pain score for use in Danish children with CP.

2. Methods

The study was approved by the local ethics committee (M-20100189) and the Danish Data protection Agency (1-16-02-97-10). Oral and written informed consent was obtained from the parents of the children, and the study was carried out in accordance with the principles of the Helsinki Declarations. The clinical validation of the r-FLACC was undertaken during 2010–13. The translation of the r-FLACC complied with the international guideline set up by the Translation and Cultural Adaptation Group (TCA) – Principles of Good Practice (PGP) [21]. The COSMIN checklist was used as a guideline in the reliability and validity testing of the r-FLACC score. The checklist is designed for HR-PROs with high complexity and since the r-FLACC score is not patient-reported some discrepancies is present resulting in the COSMIN checklist being a guideline rather than an absolute method.

Twenty-seven children were included after informed consent was obtained. Inclusion criteria were children with CP who were not able to self-assess pain and planned orthopaedic surgery at the Department of Children's Orthopedics, Aarhus University Hospital. The surgical procedure varied in severity from minor tendon surgery to major bony orthopaedic surgery.

Two methods for postoperative pain assessment were used at the same time in the postoperative hospital stay. The Observational Visual Analogue Score (VAS-OBS, range 0–10) was used at the bedside where the parents or primary caregivers estimated the pain intensity in their child during a 2-min period. A 2-min standardized video recording of the child was made in order to score the patient using the r-FLACC. In addition, in 10 of the included children, a preoperative baseline video-recording was made in order to assess construct validity of the r-FLACC score. In all the video recordings a close-up of the face, the legs and the total body of the child was performed. The parents or primary caregivers were

asked to try to console the child if they found it necessary. Furthermore, the parents or primary caregivers preoperatively completed a questionnaire regarding the individual pain behaviours for their child related to each subgroup of the r-FLACC score.

Inter-rater reliability of the r-FLACC score was tested by review of the video recording by two registered nurses (RN) experienced in the care and pain management of children with CP. They independently assigned an r-FLACC score, using the translated version of the r-FLACC, after having familiarized themselves with the parent's questionnaire of specific manifestations of pain for the child in question. Ten of the recordings were reviewed again 1 year later by one of the nurses in order to assess intra-rater reliability.

2.1. Statistical analysis

No standards for sample size calculations exists for reliability and validity testing of a measurement instrument such as the r-FLACC, hence the recommendation from Terwee et al. [22] are used. A factor analysis requires 4–10 subjects per item in the score and with the r-FLACC having 5 items we chose a sample size of 25 children. Final sample size was 27 children adjusting for possible drop-outs.

Data analysis was conducted using STATA for windows version 11. Before the statistical analysis all continuous data were plotted to assess normality. Analysis of the r-FLACC scores and VAS-OBS scores were accomplished with use of the Student's *t*-test for normal distributed data, factor analysis, Cronbachs alpha (CA), a one-way analysis of variance intra-class correlations – type 1 (ICC) [23], Bland–Altman plots and Pearson's correlation coefficients. Factor analysis uniqueness (FAU) < 0.6 was interpreted as indicating unidimensionality. ICC ranges from 0 to 1 and was interpreted using the following criteria: 0.00–0.39 poor, 0.40–0.59 fair, 0.60–0.74 good and 0.75–1.00 excellent. A Pearson's correlation > 0.70 indicates criterion validity. A *p*-value of < 0.05 was considered to be significant.

3. Results

The 27 children included were 3–15 years old, 11 girls and 16 boys with Gross Motor Function Classification Scores (GMFCS) scores ranging from II to IV. They all had tetraplegic CP and were not able to self-report pain. Twenty children underwent pelvic and femoral osteotomy, 4 children underwent tendon or soft tissue surgery, 1 child underwent calcaneal osteotomy, 1 child underwent epiphysiodesis of the distal femur and 1 child had femoral plates removed. The Danish version of the r-FLACC score was used for pain assessment. The original English version is presented in Table 1.

Reliability contains three measurement properties. Internal consistency was found to be excellent with a CA for the two raters of 0.9023 and 0.9758, respectively. The CA values for each subgroup are listed in Table 2. A factor analysis of the 5 subgroups showed low FAU indicating unidimensionality (FAU face: 0.55; FAU legs: 0.25; FAU activity: 0.31; FAU cry: 0.25; FAU consolability: 0.19).

Excellent intra-rater reliability was proved by a test–retest resulting in an ICC of 0.97530. Good inter-rater reliability between rater 1 and 2 was seen with an ICC of 0.74576. Both intra and inter-rater reliability are illustrated by Bland–Altman plots where the agreement is visualized (Figs. 1 and 2).

No floor or ceiling effects were observed and the entire range of the r-FLACC score was used with a minimum r-FLACC score of zero in 30% of the children and a maximum score of ten in 7% of the children.

Validity contains three measurement properties. Content validity was tested by the originators of the r-FLACC. Construct validity was supported by a significant increase in r-FLACC scores following

Table 1

The r-FLACC score for pain assessment in children with CP. Revisions from the original FLACC score to this r-FLACC score are noted in italics.

	r-FLACC score for pain assessment in children with cerebral palsy
Face	
0 point	No particular expression or smile
1 point	Occasional grimace/frown; withdrawn or disinterested; <i>appears sad or worried</i>
2 point	Consistent grimace or frown; frequent/constant quivering chin, clenched jaw; <i>distressed-looking face; expression of fright or panic</i>
<i>Individualized behaviour</i>	
Legs	
0 point	Normal position or relaxed; <i>usual tone & motion to limbs</i>
1 point	Uneasy, restless, tense; <i>occasional tremors</i>
2 point	Kicking, or legs drawn up; <i>marked increase in spasticity, constant tremors or jerking</i>
<i>Individualized behaviour</i>	
Activity	
0 point	Lying quietly, normal position, moves easily; <i>regular, rhythmic respirations</i>
1 point	Squirming, shifting back and forth, <i>tense or guarded movements; mildly agitated (e.g. head back and forth, aggression); shallow, splinting respirations, intermittent sighs</i>
2 point	Arched, rigid or jerking; <i>severe agitation; head banging; shivering (not rigours); breath holding, gasping or sharp intake of breaths, severe splinting</i>
<i>Individualized behaviour</i>	
Cry	
0 point	No cry/verbalization
1 point	Moans or whimpers; occasional complaint; <i>occasional verbal outburst or grunt</i>
2 point	Crying steadily, screams or sobs, frequent complaints; <i>repeated outbursts, constant grunting</i>
<i>Individualized behaviour</i>	
Consolability	
0 point	Content and relaxed
1 point	Reassured by occasional touching, hugging or being talked to. Distractible
2 point	Difficult to console or comfort; <i>pushing away caregiver, resisting care or comfort measures</i>
<i>Individualized behaviour</i>	

Table 2

Internal consistency illustrated by Cronbachs alpha (CA) for both raters for each of the 5 subgroups in the r-FLACC pain score.

	CA rater 1 (n = 27)	CA rater 2 (n = 20)
Face	0.9147	0.9756
Legs	0.8777	0.9729
Activity	0.8763	0.9744
Cry	0.8697	0.9630
Consolability	0.8618	0.9630
Total	0.9023	0.9758

surgery compared to preoperative assessments (*n* = 17; difference 2.23; *p* = 0.0397). Criterion validity was established only in rater 1 by a Pearson's correlation coefficient of 0.76 (*n* = 27) with rater 1 and 0.59 (*n* = 20) with rater 2 when comparing the r-FLACC scores of the two raters and the VAS-OBS scores.

4. Discussion

The current study establish excellent validity and reliability of the Danish version of the r-FLACC pain score for assessing

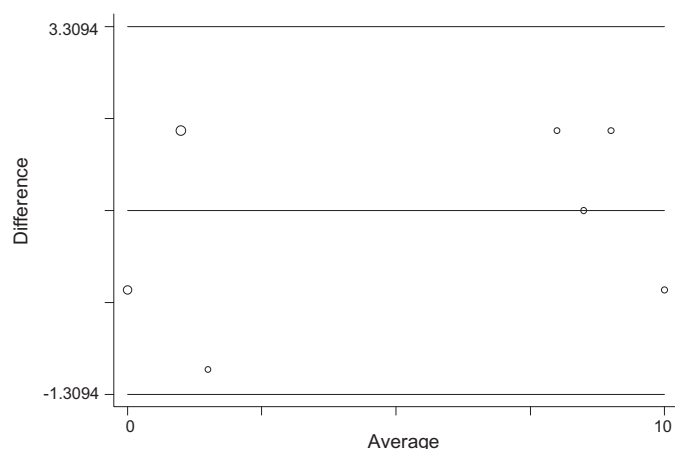


Fig. 1. Agreement (intra-rater reliability) illustrated by Bland–Altman plot ($n=10$) with comparison between the first r-FLACC scores from rater 1 and the re-test scores after 1 year. The differences between the pairs of measurements on the vertical axis are plotted against the average of each pair on the horizontal axis. The middle horizontal line reflects the mean difference (1.0 (95% CI: 1.074; 1.826)) and the upper and lower line the limits of agreement (–1.309; 3.309).

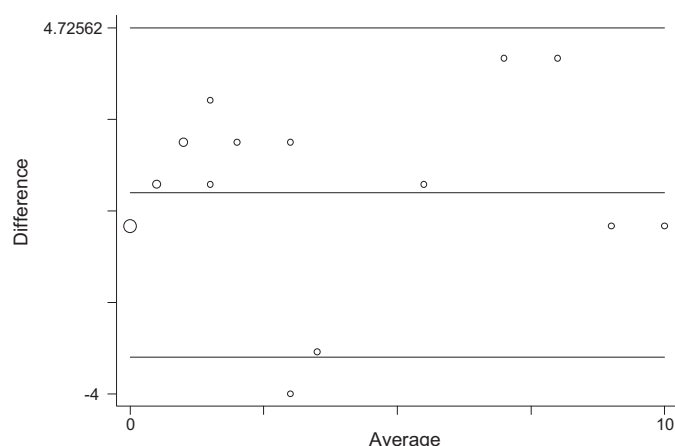


Fig. 2. Agreement (inter-rater reliability) illustrated by Bland–Altman plot ($n=20$) with comparison between rater 1 and rater 2. The differences between the pairs of measurements on the vertical axis are plotted against the average of each pair on the horizontal axis. The middle horizontal line reflects the mean difference (0.8 (95% CI: –0.119; 1.719)) and the upper and lower line the limits of agreement (–3.126; 4.726).

postoperative pain in children with CP not able to self-report pain. In comparison to the original r-FLACC validation study, this study benefits from a homogenous study population only including cognitively impaired children, who were not able to self-report pain and by using the COSMIN checklist as a guideline. On the other hand, the study by Malviya et al. [4] benefits from a larger study population and includes both bedside and video assessments.

The internal consistency was found to be excellent for both raters but also for each item of the score assessed by both CA and a factor analysis confirming that all 5 items of the score are sufficiently correlated and measures the same construct, hence unidimensionality is established. However, Scholtes et al. [17] points out that internal consistency is established if CA values are between 0.70 and 0.95 and if above 0.95 the measurement instrument may contain too many items that are assessing the same underlying construct. This study finds a CA of 0.90 and 0.98 (two raters), which supports internal consistency. Since the r-FLACC is designed to measure pain as a very specific construct the high CA is expected and further establishes unidimensionality. This is opposite to other PROMs measuring more unspecific constructs, for example quality of life. The subgroup CA range from 0.86 to 0.98 with consolability

having the lowest value for both rater 1 and 2, indicating that even though a high internal consistency, consolability is the item that least fit the score. The COSMIN checklist is developed for Health-Related Patient-Reported Outcomes (HR-PROs) often characterized by more items with higher complexity than in the r-FLACC pain score. Furthermore the r-FLACC is not a patient-reported outcome, but rather a tool for the clinician in measuring pain. Considering these issues a higher CA value than referenced for classical HR-PROs would be expected for the r-FLACC.

The intra-rater reliability in present study was excellent (ICC: 0.97530); but the inter-rater reliability only good (ICC: 0.74576). The Bland–Altman plots for both reliability measures illustrates that the differences of measurements are concentrated around the mean difference, though with a rather large interval of agreement for especially inter-rater reliability. This is in accordance to other studies testing the psychometric properties of pain assessment scales, where the range of ICCs for inter-rater reliability for the INRS was 0.65–0.87 [10], for the NCCPC-PV 0.78–0.82 [11] and for the PPP 0.74–0.89 [13]. This suggests that even though intra-rater reliability is established, the most exact pain assessment during hospitalization of a child with CP is performed using only one rater continuously for each child.

When developing an instrument as for example the r-FLACC all known relevant pain variables for this patient group have to be included in the instrument, hence content validity is a subjective assessment of the degree to which the r-FLACC adequately reflects the level of pain. The assessment of content validity of the r-FLACC score has been described by Malviya et al. [4] by a thorough review of pain behaviours common to individuals with cognitive impairment resulting in an expansion of descriptors in the least reliable categories of the original FLACC score (Table 1). This establishes content validity of the r-FLACC score and is transferable through translations.

In present study, construct validity was demonstrated by an increase in the level of pain after surgery. By definition construct validity estimates the degree to which the scores of the measurement instrument are consistent with predefined hypotheses on internal and external relationships [17]. In the case of pain assessment, it is a predefined hypothesis that the level of pain will increase following surgery showing an expected difference. Construct validity was also demonstrated by Malviya et al. [4] after revision of the original FLACC score by demonstration of a decrease in pain scores after administration of analgesics. The discussion between using agreement or reproducibility parameters is pointed out by de Vet et al. [24] and for a pain score like the r-FLACC agreement parameters could be more illustrative than reproducibility parameters. Nevertheless, in medical science, reproducibility parameters are more commonly used. Present study opted to calculate and report reproducibility by use of ICC similar to Solodiuk et al. [10] and in accordance to the COSMIN guideline. Opposite, Malviya et al. [4] calculates both kappa agreement and ICC reproducibility, making comparisons between the two studies difficult.

To assess criterion validity a comparison to another gold standard instrument with known good validity is a prerequisite. For the assessment of pain in children several instruments exists in the Scandinavian countries but none serves as a gold standard instrument. The Visual Analogue Score has been validated for use in several patient groups and is the mostly used pain score in Denmark and since no perfect gold standard for the evaluation of pain in children with CP exists, the VAS-OBS was chosen as an approximation for a gold standard instrument, knowing that it may be a limitation to the study. The r-FLACC did demonstrate acceptable criterion validity for rater 1 but not for rater 2, which may be caused by a number of factors. Firstly, Scholtes et al. [17] states that if the comparison instrument is not ‘really gold’, one would not know,

which instrument is not valid if the comparison between the two instruments is low. Secondly, the VAS-OBS and r-FLACC scores have fundamental differences that might interfere with the comparison of the scores. One difference is that VAS-OBS involves assessment by parents and the r-FLACC involves assessment by a RN. Previous studies have found that parents tend to overestimate pain in comparison to other observers [4]. In addition, the r-FLACC score is assessed by use of a video-recording and the VAS-OBS score is assessed bedside. These factors might influence the comparison of the two instruments. In the study by Malviya et al. [4] criterion validity was determined by comparison to both the Nursing Assessment of Pain Intensity (NAPI), developed for healthy older infants and children [25] and for the children able to self-report by the Verbal 0–10 Numbers Scale or Simplified Faces Scale [4]. Some limitations are noted due to the NAPI not being a gold standard for pain assessment in children with CP and important demographic differences exists by comparing r-FLACC scores from children not able to self-report to a number or face scale from children self-reporting.

Assessment of r-FLACC scores *via* video-recordings is an advantage when estimating reliability, due to the fact that the level of pain of the child on the video-tape is unchanged and actual reliability estimates the extent to which scores for a patient who have not changed are the same for repeated measurements. The long interval is a benefit since it prevents recollection by the rater and this far outweighs the disadvantage that a possible change in the way the rater evaluates pain might occur.

The validation process was designed only to include children with postoperative pain and not procedural or chronic pain hence possibly limiting the generalizability of the r-FLACC pain score. However, by using a postoperative setting the risk of floor or ceiling effects was avoided since the entire range of the measurement instrument has been applied. In addition, no floor or ceiling effects were observed indicating that this measurement instrument is valid in measuring no, average and severe pain in children with CP.

5. Conclusions

This study benefits from a systematical approach to the validation and reliability parameters by using the COSMIN checklist as a guideline. Results showed excellent internal consistency, excellent intra-rater reliability and a good inter-rater reliability. In addition, excellent construct validity and criterion validity were established, though only for rater 1. Since a validation study of the original language revealed similar results it is evident that the r-FLACC pain score maintains its psychometric properties after translation. In conclusion, the r-FLACC pain score is both valid and reliable in assessing postoperative pain in children with CP not able to self-report pain.

6. Implications

With the r-FLACC clinicians have a valid tool for assessing the level of postoperative pain, hence increasing the quality of pain management in children with CP. In addition the valid r-FLACC score has the potential for use in interventional research regarding pain management in this vulnerable group of patients. The r-FLACC is valid and reliable in measuring postoperative pain and future perspectives include validation of the r-FLACC score for procedural and chronic everyday pain and implementation into daily practice in the Scandinavian countries.

Conflicts of interest

The authors have no conflict of interest.

Acknowledgements

The authors acknowledge Lene E. Sloth and Lene Ahrensbach for the meticulous revision of the video recordings.

The study was supported by a grant from the Elsass Foundation.

References

- [1] Malviya S, Voepel-Lewis T, Tait AR, Merkel S, Lauer A, Munro H, Farley F. Pain management in children with and without cognitive impairment following spine fusion surgery. *Paediatr Anaesth* 2001;11:453–8.
- [2] Penner M, Xie WY, Binopal N, Switzer L, Fehlings D. Characteristics of pain in children and youth with cerebral palsy. *Pediatrics* 2013;132:e407–13.
- [3] Parkinson KN, Dickinson HO, Arnaud C, Lyons A, Colver A, on behalf of the SPARCLE group. Pain in young people aged 13 to 17 years with cerebral palsy: cross-sectional, multicentre European study. *Arch Dis Child* 2013;98:434–40.
- [4] Malviya S, Voepel-Lewis T, Burke C, Merkel S, Tait A. The revised FLACC observational pain tool: improved reliability and validity for pain assessment in children with cognitive impairment. *Paediatr Anaesth* 2006;16:258–65.
- [5] Voepel-Lewis T, Malviya S, Tait AR, Merkel S, Foster R, Krane EJ. A comparison of the clinical utility of pain assessment tools for children with cognitive impairment. *Paediatr Anaesth* 2008;106:72–8.
- [6] Chen-Lim ML, Zarnowsky C, Green R, Shaffer S, Holtzer B, Ely E. Optimizing the assessment of pain in children who are cognitively impaired through the quality improvement process. *J Pediatr Nurs* 2012;27:750–9.
- [7] McGrath PJ, Rosmus C, Camfield C, Campbell MA, Hennigar A. Behaviours caregivers use to determine pain in non-verbal, cognitively impaired children. *Dev Med Child Neurol* 1998;40:340–3.
- [8] Drendel AL. Pain assessment for children overcoming challenges and optimizing care. *Pediatr Emerg Care* 2011;27:773–81.
- [9] Voepel-Lewis T, Merkel S, Tait AR, Trzcinka A, Malviya S. The reliability and validity of the Face, Legs, Activity, Cry, Consolability observational tool as a measure of pain in children with cognitive impairment. *Anesth Analg* 2002;95:1224–9.
- [10] Solodiuk JC, Scott-Sunderland J, Meyers M, Myette B, Shusterman C, Karian VE, Harris SK, Curley MAQ. Validation of the Individualized Numeric Rating Scale (INRS): a pain assessment tool for nonverbal children with intellectual disability. *Pain* 2010;150:231–6.
- [11] Breau LM, Finley GA, McGrath PJ, Camfield CS. Validation of the Non-communicating Children's Pain Checklist-Postoperative Version. *Anesthesiology* 2002;96:528–35.
- [12] Johansson M, Carlberg EB, Jylli L. Validity and reliability of a Swedish version of the Non-Communicating Children's Pain Checklist-Postoperative Version. *Acta Paediatr* 2010;99:929–33.
- [13] Hunt A, Goldman A, Seers K, Crichton Moffat V, Oulton K. Clinical validation of the Paediatric Pain Profile. *Dev Med Child Neurol* 2004;46:9–18.
- [14] Collignon P, Giusiano B. Validation of a pain evaluation scale for patients with severe cerebral palsy. *Eur J Pain* 2001;5:433–42.
- [15] Crosta QR, Ward TM, Walker AJ, Peters LM. A review of pain measures for hospitalized children with cognitive impairment. *Pediatr Nurs* 2013;19:109–18.
- [16] Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, de Vet HCW. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010;10:22.
- [17] Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable. *Injury* 2011;42:236–40.
- [18] Crellin D, Sullivan TP, Babi FE, O'Sullivan R, Hutchinson A. Analysis of the validation of existing behavioral pain and distress scales for use in the procedural setting. *Paediatr Anaesth* 2007;17:720–33.
- [19] Ailliet L, Rubinstein SM, de Vet HCW, Tulder MW, Terwee CB. Reliability, responsiveness and interpretability of the neck disability index-Dutch version in primary care. *Eur Spine J* 2014;24:88–93.
- [20] Talma H, Chinapaw MJM, Bakker B, HiraSing RA, Terwee CB, Altenburg TM. Bioelectrical impedance analysis to estimate body composition in children and adolescents: a systematic review and evidence appraisal of validity, responsiveness, reliability and measurement error. *Obes Rev* 2013;14:895–905.
- [21] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8:94–104.
- [22] Terwee CB, Bot SDM, Boer MR, Windt DAVM, Knol DL, Dekker J, Bouter LM, Vet HCW. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34.
- [23] McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30–46.
- [24] de Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006;59:1033–9.
- [25] Schade JG, Joyce BA, Gerkenmeyer J, Keck JF. Comparison of three preverbal scales for postoperative pain assessment in a diverse pediatric sample. *J Pain Symptom Manage* 1996;12:348–59.