**Appendix A Psychometric evaluation studies: Study details**

**Systematic Review of the FLACC scale for assessing pain in infants and children: is it reliable, valid, useful and feasible for use?**

Dianne CRELLINa-c, Nick SANTAMARIAa f Franz E BABLb-d, Denise HARRISONa b e

|  |  |  | **Results (inc quality score)** | | |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Aim/Design and method** | **Subjects/circumstances/ Setting/Pain measures** | **Reliability** | **Validity** | **Feasibility & clinical utility** | **Comments** |
| Original study | | | | | | |
| Merkel et al, 1997 [[13](#_ENREF_13)] | To evaluate reliability and validity of FLACC Tool  Descriptive repeated measures study  Phase 1 – 2 independent assessors score with FLACC 3 times at 5min intervals. Bedside nurse also score at final time point using global scale.  Phase 2 - assessment before & after analgesic  Phase 3 - FLACC and OPS applied by 2 assessors blinded to each other’s scores. | 89 children postop aged 2mth – 7 years (mean = 3.0 ±2.0)  Phase 1 – 30 children  Phase 2 – 29 children  Phase 3 – 30 children  Pain: Post-operative  Setting: PACU  Index: FLACC  Reference: OPS | Inter-rater: Correlation between observers r[87]=0.94, p<.001) Kappa values for items range from 0.52 (face) – 0.82 (cry) COSMIN - poor | Content: behaviours selected that had been described and validated in other tools (eg: CHEOPS, OPS TPPPS and Buttner/Finke). Piloted and revisions made.  Hypothesis (convergent): Positive correlation between FLACC and OPS (r=.80, p<.001) COSMIN - poor  Responsiveness: FLACC scores decreased post-analgesic from pre=7.0 ± 2.9 to 10min=1.7 ±2.2, 30min=1.0 ±1.9, 60min=.02 ±.05 (p<0.001 at each interval). COSMIN - poor | Not assessed | Designed to offer more feasible scale (shorter, more easily remembered). FLACC not obviously shorter (5 v 3-6 items) and feasibility not tested  FLACC comprised of items from existing scales (OPS, Buttner/Finke< CHEOPS etc) – validation included correlating with exisiting scales (positive results predictable). |
| FLACC repeat validation studies (n = 3) | | | | | | |
| Bringuier et al, 1999 [[3](#_ENREF_3)] | To compare the psychometric properties, sensitivity and specificity of CHEOPS, CHIPPS, FLACC and OPS (collectively BRS)  Comparative longitudinal prospective study  Children videoed for scoring at 4 time points  T1 – day before surgery  T2 – pre-induction  T3 – PACU 20min post extubation  T4 – morning post surgery  4 raters scored each video using each scale in random order. Children (> 4years old) or parents rated pain on FPS-R and anxiety on VAS-anxiety scale.  Group of nurses assessed face validity of each scale  FASS used to establish criterion index to evaluate validity of scales. | 148 children generating 511 videos for children mean age 2.9 years (range 1 – 7 years)  Setting: inpatient surgical centre, France  Pain: Post-operative  Index: CHEOPS, CHIPPS and OPS  Reference: na | Inter-rater: ICC observers was >0.86.  COSMIN - good  Internal consistency – Cronbach’s alphas ranged from 0.81 to 0.93. Cronbach’s alpha for CHEOPS (0.81) higher without 2 items – complaint & touch (0.83; 0.82)  COSMIN – good  FLACC results not described separately | Content (face): FLACC and CHIPPS accepted by experts. Scoring out of 10 with cutoff of 3 preferred. COSMIN - poor  Structural (construct): principle component analysis showed that FLACC, CHIPS and OPS were homogeneous. All item correlations >0.4, the two lowest items from CHEOPS (r=.48 complaints and touching wound).  COSMIN - excellent  Hypothesis (convergent): correlations between the 4 scales were 0.88–0.94. Correlation between the 4 scales & self-reports of pain only significant at T3 and T4. (OPS at T4, p >0.5). Correlation between BRS & FASS 0.71 – 0.78 (p<0.5)  Correlation between FLACC and FACES scores (r(30) = 0.584, p=0.001). FLACC did not correlate with scores for children aged < 5years (r(14) = 0.254, p = 0.381). For children aged >5y 9r[16] = 0.830; p=0.0001).  Hypothesis (discriminant): – correlations with anxiety only significant at T2 when anxiety assessed by parents (0.23–0.34) at T3 and T4 when anxiety assessed by child (T 3 : 0.63–0.77; T4: 0.54–0.78) and parents (T3: 0.22–0.25; T4: 0.27–0.37). Correlation coefficients higher using self-reports (t3: 0.63–0.77; t4: 0.54–0.78) than proxy reports of anxiety (T3: 0.22–0.25; T4: 0.27–0.37).  COSMIN - fair  Responsiveness: All scales changed over time (p<.001). CHEOPS item – ‘touched the wound’ rarely seen. COSMIN - fair | Utility: discrimination (pain versus no pain) Specificity - FASS as reference: = 96% and FPS-R as reference = 89%. Sensitivity – FASS = 77% and FPS-R = 62).  Risk factor for false negatives - silence (OR adjusted = 4.47, 95%CI: 1.71 – 11.55) and for false positives - level of parental-reported anxiety (p=.04) | Scoring of multiple tools may impact on convergence  Only 32% of children provided self report – numbers not increased in older aged children  BRS did not rate pain pre-op as 0 – authors conclude restlessness contributes to false positives  High correlation with anxiety but did not increase number of false positives  Only 11 children able to report their anxiety in PACU .  FLACC high sensitivity and highest specificity of the 4 scales. However, more likely to result in false negative than false positive  Pain under-reported – silence likely confounder - contributing to false negatives.  Potential that all scale items can’t be adequately assessed from video footage. |
| Willis et al, 2003 [[24](#_ENREF_24)] | To further test the validity of the Faces, Legs, Activity, Cry and Consolability (FLACC) Behavioural Pain Assessment Scale  Descriptive observational study  Pain was scored post-operatively by nurse researcher using FLACC. Children independently self-reported pain using the FACES scale. 2nd nurse simultaneously & independently scored using FLACC. | 30 children aged 3 – 7 years (5.01 ± 1.04)  Setting: inpatient units  Pain: Post-operative  Index: FLACC  Reference: Self-report - Faces Scale | Inter-rater agreement = 100% for 6 paired observations (17% of observations) COSMIN – poor | Criterion (concurrent): Correlation between FLACC and FACES scores (r(30) = 0.584, p=0.001). FLACC did not correlate with scores for children aged < 5years (r(14) = 0.254, p = 0.381). For children aged >5y 9r[16] = 0.830; p=0.0001) COSMIN - fair | Not assessed | Children 3-5y unable to adequately use faces scale most likely explanation  Research team includes members of development and original research team |
| FLACC validation for alternate circumstances (age, pain, language) (n = 15) | | | | | | |
| Ahn et al, 2007 [[1](#_ENREF_1)] | To examine pain-like responses to frequent stimulants in the neonatal intensive care unit (NICU) using CRIES, FLACC and PIPP, and determine the clinical feasibility and validity of these tools  Exploratory correlational study  Observations of baseline prior to and 8 different stimuli categorised as:  A - invasive  B - routine care  C - auditory stimulants  made by researcher using all three scales.  Multiple observations from for each infant possible | Sample: 110 consecutively enrolled infants mean age GA 32.43 weeks at birth – testing at 1 week of age  \* Sedated infants and those with congenital & neurological anomalies excluded  274 observations made across Groups A, B and C.  Setting: NICU  Pain: Procedural  Index: FLACC, CRIES and PIPP  Reference: na | Inter-rater: assessed using 10 cases BEFORE data collection – results not reported | Hypothesis (known groups): Significant hierarchy for mean scores of the 3 groups for CRIES (F(2,271) =125.285, p<.001), FLACC ( F(2, 271)=88.257, p<.001) and PIPP ( F(2,271) =56.504, p<.001). Group A highest mean pain scores for all three tools (p <.01) .  Hypothesis (convergent): Strong correlation between CRIES and FLACC in each category (r =.826, .843, and .824 for A, B and C, respectively; p<.01 in all). Low correlation between PIPP and CRIES and FLACC, although all 3 measures were significantly related (.292<r<.521, p<.01)  Pain scores higher in full-term infants than in premature infants using CRIES (2.78 v 1.95; p<.001) and FLACC (2.52 v 1.72; p<.01). Mean PIPP score from group C was lower in full-term infants than in premature infants (3.10 v 4.28; p<.01)  COSMIN - fair | Not tested | Scales applied randomly by single assesor except PIPP (last as required 30 sec delay to apply correctly – may have impacted on lower correlations between PIPP and CRIES and FLACCs.  Scales all differentiated between the different levels of care. However, routine care associated with elevated scores – therefore painful or scales measuring another construct.  Age related differences imply inadequacy of FLACC and CRIES for preterm infants. Superiority of PIPP claimed on the basis of higher scores for preterm experiencing auditory stimulus – however, auditory stimulus not painful. |
| Bai et al, 2012 [[2](#_ENREF_2)] | To identify 1) concurrent validity of the FLACC and COMFORT-B scales for pain assessment in Chinese children after cardiac surgery; 2) to evaluate the sensitivity, specificity, and the optimal FLACC and COMFORT-B scale cutoff scores; and 3) to explore factors that predict COMFORT-B and FLACC scores  Repeated observation study  VASObs, FLACC and COMFORT-B measures taken 2hrly during the day on day 0, 1 and 2 post-op – total of 18 measures  FLACC and COMFORT-B translated into Chinese. Content validity of COMFORT-B (chinese) tested using 3 experts  Testing at various cut-offs for FLACC and COMFORT-B to determine sensitivity and specificity for detecting pain / no pain (defined by expert applied VASObs (<4 = not in pain)  Multiple regression analysis to determine predictors. | 174 children aged 0 – 7 (median 8 months)  (4 excluded – data for 170)  Setting: CICU, China\*  Pain: Post-operative (cardiac surgery)  Index: FLACC, COMFORT B scale (Chinese)  Reference: VASobs | Inter-rater: testing results from assessment PRIOR to data collection reported – 4 assessments undertaken by two researchers – intra-class correlation FLACC = 0.84, COMFORT-B = 0.98 | Criterion (concurrent): VASobs high correlation with FLACC (r =0.86; p= .0001) & low correlation with COMFORT-BChinese (r=0.31; p=.0001). COMFORT-BChinese score moderately correlated with FLACC (r=0.51; p=.0001). COSMIN - poor  Hypothesis (convergent) No correlation btw scores and physiological markers (HR, ArtBP) p>.05  Multiple regression: FLACC higher scores assoc with younger age (p<.001) & relaxants (p=.021). Higher COMFORT-B scores assoc with decreased duration ventilation (p<.001) & lower age (p=.028), Lower scores assoc with analgesics (p=.008) & relaxants (p=.025).  COSMIN - fair | Utility: COMFORT-B and FLACC scores for children in pain (VASobs≥4) were significantly higher than scores for children not in pain [VASobs<4] (p<0.0001).  Used to establish cutoff – FLACC ≥2 sensitivity 98% and specificity 88% COMFORT-BChinese cutoff ≥13 sensitivity = 86% and specificity = 83% | VASobs used as reference scale. However, Van Dijk, 2002 – cites correlation with self report - 0.23 - .83, therefore questionable choice for reference scale.  Correlations reported as criterionvalidity - observational scale used as reference – not a gold standard  Impact of medications not addressed – observations made following muscles relaxants or sedation in many cases - may impact on behaviour and therefore scores. Aim for haemodynamic stability, children receiving haemodynamically active medications (not reported) therefore unable to determine impact on physiological markers.  Lower cut-off for pain (≤2) than shown previously for FLACC (may reflect the population - sedated) |
| Da Silva et al, 2008 [[6](#_ENREF_6)] | To translate, back-translate and cross-culturally adapt the content of the FLACC (Face, Legs, Activity,  Cry, Consolability) and Faces Pain Scale-Revised (FPS-R) scales for the evaluation of pain in Brazilian young students and adolescents.  Three stage design  Translation and back translation from English to Brazilian Portuguese  Survey of 12 expert health professionals to assess cross cultural adaptation and content  Pretest: FLACC – survey of clinicians to assess ability to understand & apply the scale. FPS-R – survey of children about their ability to use scale | 20 oncology patients aged 7 – 17 years  Setting: outpatients and in patient ward in Brazilian  22 health professionals  Index: FLACC (Brazilian Portuguese) | Not assessed | (Cross cultural) Face and content – Changes were made to the Brazilian translations from the literal translation to one where the intention was better expressed  Mean score (scale 0 - 10) for comprehension of the FLACC scale was 9.6 (±1.0).  COSMIN - poor | Not assessed | Full breadth of cross cultural validity assessment not completed  Assessor comments acknowledged some ambiguity in the descriptors for scoring - amendments made to scale to suit Brazilian application |
| da Silva et al, 2011 [[5](#_ENREF_5)] | The aim of this research is to examine the validity and reliability of the Brazilian version of the Revised Faces Pain Scale and the Face, Legs, Activity, Cry, Consolability scale.  Prospective observational validation study  Children with canacer diagnosis rated pain using FPS-R and simultaneously physician applied FLACC | 90 children aged 7 – 17 years  Setting: inpatient and outpatient  Pain: Secondary to disease (oncology)  Index: FLACC, Revised Faces Pain Scale (Brazilian) | Internal consistency: cronbach’s α – 0.76, correlations between items ranged from 0.12 – 0.65.  COSMIN - fair | Criterion (concurrent): Spearman’s correlation between FLACC and FPS-R = 0.74.  Mean FPS-R score 1.74 (SD 2.43), mean FLACC score = 0.78 (SD 1.44)  COSMIN - good | Not assessed |  |
| Gomez et al, 2013 [[8](#_ENREF_8)] | To establish inter-rater and intrarater agreement of the FLACC scale in toddlers during immunization.  Observational valiation study  Children videotaped during immunisation procedure (Two raters scored video segments in random order and one set of raters rescored video segments 3 weeks later). FLACC scored at 4 time points, prior to immunisation, during insertion of needle and 15 and 30 seconds following completion of immunisation | 30 children aged 12 – 18 months  Setting: Immunisation drop in service  Pain: Procedural (immunisation)  Index: FLACC  Reference: not applicable | Intra-rater: ICC were 0.88 at baseline, 0.97 at insertion of first needle, and 0.80 & 0.81 at 15 s and 30 s following the final injection, respectively.  Inter-rater: ICC were 0.40 at baseline, 0.95 at insertion of first needle, and 0.81 and 0.78 at 15 s and 30 s following the final injection, respectively.  COSMIN - good | Not assessed | Not assessed | Raters blinded to each other and time delay and random order of presentation of video segments designed to reduce memory of segments for second application of FLACC for intra-rater reliability.  Able to view video segment multiple times before scoring - may alter reliability results impacting on capacity to generalise to practice |
| Johansson et al, 2009 [[10](#_ENREF_10)] | To evaluate the concurrent validity and reliability of Swedish versions of the behavioural COMFORT and a modified version of the FLACC scale for assessment of pain and sedation in intubated and ventilated children and to evaluate the construct validity of the FLACC scale for assessment of pain.  Prospective observational study  6 nurses trained to use scales, piloted to establish acceptable agreement  40 children - 2 out of the 6 nurses applied both scales in random order at random times of day and 2 bedside nurses assessed using VASobs & NIS score  Another 20 children – 1 nurse assessed FLACC scores before and after analgesic.  Scales translated into Swedish using forwards and backwards method | 40 children aged 0 – 108 months (median 4 months) resulting in 119 paired observations  20 additional children aged 1 – 13 months (median 4months)  Setting: PICU  Sweden  Pain: Postoperative (cardiac)  Index: FLACC (modified item - cry, Swedish), COMFORT scale (Swedish)  Reference: VASobs, Nurse interpretation of Sedation (NIS) | Inter-rater: weighted kappa scores for FLACC scores 0.63 (95% CI 0.53–0.72) and COMFORT-B scores 0.71 (95% CI = CI 0.65–0.77)  Weighted kappa for individual items for FLACC varied from 0.51 (activity) – 0.61 (face).  COSMIN – good | Criterion (concurrent) – Correlations between FLACC and VASobs 0.50 (p <0.05), FLACC and NIS 0.50 (p <0.05), COMFORT-B and VASobs,= 0.49 (p <0.05) and COMFORT-B and NIS 0.57 (p <0.05) Correlation between COMFORT-B and FLACC = 0.76 (p <0.05). COSMIN - poor  Responsiveness – median FLACC score decreased from 5 to 0–2 (p <0.001, Wilcoxon signed rank test) following morphine.  COSMIN - poor | Utility: median FLACC score for VASos <3 = 0.5 (0 – 10) and VASobs>3 = 3.5 (0-8) and median COMFORT-B scores VASos <3 = 12 (6 - 21) and VASobs>3 = 17 (11-23) (Kruskal-Wallis, p <0.01).  FLACC scores for three levels of sedation were 0 (0–3) = ‘oversedated’, 0 (0–8) = ‘adequately sedated’ and 4 (0–8) = insufficiently sedated’ (Kruskal-Wallis, p<0.01). COMFORT-B scores for the 3 levels of sedation were 9 (6–15), 12 (6–21) and 16 (7–23) respectively (Kruskal-Wallis, p <0.001). | VASobs used as reference scale. However, Van Dijk, 2002 – cites correlation with self report - 0.23 - .83. Therefore questionable choice for reference scale  Correlations reported as criterionvalidity - observational scale used as reference – not a gold standard  Scale modified for use in intubated critically ill children – therefore cry altered to ‘cry face or moaning’ – no content validation attempted  Reliability for FLACC slightly less than shown in other studies – may be result of modifications (reliability not lowest for ‘cry’)  Only 7 patients with VASobs>3 therefore data supports reliability & validity in lower pain states only.x |
| Malviya et al, 2006 [[11](#_ENREF_11)] | To revise the FLACC tool to include behaviours more specific to children with cognitive impairment (CI) and evaluate the reliability and validity of the revised FLACC (modified descriptors) for assessment of pain in children with CI  Observational repeated measures comparison study  Scale revision using behaviours common to children with CI (literature) & those seen in children with CI videoed following surgery. Content validated by experts. Parents individualised scale  FLACC (2 nurses), parental (VASobs) and child’s self-reported pain scores recorded independently post-op before & after analgesic.  Randomly ordered vi deotaped segments scored independently by 4 nurses blinded to treatment using FLACC & NAPI. 2 nurses assigned scores to 20 randomly selected segments 3-4 weeks later. | 52 cognitively impaired children aged 4 – 19 years provided 80 observations  Setting: recovery and ward  Pain: Post-operative  Index: FLACCr (modified descriptors)  Reference: VASobs | Inter-rater: ICC = 0.75 (activity) – 0.87 (cry) and total score - 0.9 (CI: 0.87 - 0.92) p< 0.001 and kappa scores 0.44 (legs) – 0.57 (face) total score 0.5.  Intra-rater: ICC = 0.97 (CI: 0.92 – 0.99)  COSMIN - good | Content (Face) – confirmed by expert physicians and advanced practice nurses  Hypothesis (convergent): Correlations between FLACC (nurse, bedside nurse and video observer) and NAPI (video observer) = 0.78 – 0.87 p<0.01, FLACC and parent VASobs = 0.65 – 0.82 p<0.01, FLACC and child report – 0.67, p = 0.051 (video observer) – 0.86, p<0.01 (bedside observer)  COSMIN - good  Responsiveness: FLACC scores decreased following analgesic assessed by both video (6.1 ± 2.6 vs 1.9 ± 2.7; p < 0.001) and bedside observers (6.1 ± 2.5 vs 2.2 ± 2.4; p < 0.001) using Wilcoxon signed rank test  COSMIN - poor | Utility: FLACC scores were coded as mild (0–3), moderate (4–6) and severe (7–10) - previously defined. Reliability for clinically relevant categories. ICC = 0.83 (CI = 0.78 – 0.86 ) | Methodology has overcome most study flaws likely to bias results.  Potential that all scale items can’t be adequately assessed from video footage.  Author a member of original scale development and validation study team |
| Manworren et al, 2003 [[12](#_ENREF_12)] | To validate the FLACC Pain Assessment Tool as a clinical tool for assessing pain and evaluating pain management interventions in preverbal children  Descriptive repeated measures comparison study  Nurses assigned FLACC score when child assessed as in need of analgesic and then at regular intervals post administration of analgesic (10min, 30min and 60min) | 147 children aged 1 day – 34 months (mean 1 yr 40 days)  Setting: PACU, PICU, surgical trauma unit and haem/onc unit,  Pain: assessed as experiencing pain by nurse (and prescribed analgesic)  Index: FLACC  Reference: na | Established prior to commencement of study 25 nurse participants assigned FLACC to 10 videos – average weighted kappa >0.54. 19 demonstrated average weighted kappa > 0.6 and included as data collectors (14 collected data) | Responsiveness: FLACC scores before [7.03 (6.66 – 7.41)] and after analgesic [30min = 2.05 (1.68 – 2.43) and 60min = 0.74 (0.48 – 1.0)] significantly different (p<0.001)  COSMIN – poor | Utility: Pre-analgesic FLACC scores significantly higher for opioid group than other analgesic groups (F[2,144]=5.55, p<0.005).  No significant difference in FLACC scores post analgesic based on analgesic group. | 5 nurses did not collect data – didn’t identify patients  Authors conclude that scale is feasible ‘easy to use’ based on speed (10min) with which nurses trained to achieve >0.54 interrater reliability.  Efforts made to include all pain levels |
| Nilsson et al, 2008 [[14](#_ENREF_14)] | To evaluate the concurrent and construct validity and the interrater reliability of the Face, Legs, Activity, Cry and Consolability (FLACC) scale during procedural pain in children aged 5–16 years.  Repeated measures study  2 observers assessed child before, during and after procedure for half sample. One assessor only for second half of sample  Children self-rated pain using CAS and distress using the FAS  Scales translated into Swedish and back translated | 80 children aged 5 – 16 (mean age 10.5)  Setting: Surgical and oncology units.  Sweden\*  Pain: Procedural pain (Cannula and percutaneous puncture to access subcutaneous access device)  Index: FLACC (Swedish)  Reference: CAS and FAS | Inter-rater: weighted kappa coefficient for total FLACC score measured during the procedure = 0.85 (p < 0.001).  COSMIN - fair | Criterion (concurrent): Spearman correlation between FLACC scores and self report CAS (r = 0.59, P < 0.05, 5–10y = 0.59 and 11-16y = 0.5) and FAS (r = 0.35, P < 0.05). COSMIN - fair  Responsiveness: median FLACC scores before increased from 0 to 1 during (p<0.0001) and decreased to 0 after the procedure (P < 0.001) Wilcoxon rank sum test. CAS Scores before during and after were 0, 0.75 and 0 respectively and FAS scores were 0.37, 0.47 and 0.37 respectively. COSMIN – fair | Not assessed | Patient group with chronic illness therefore increased pain experience likely, potentially altering pain experience and behavioural expression  Very low scores in study therefore doesn’t contribute to assessment of validity of scale to measure moderate and severe pain.  Pain scores very low (during procedure median FLACC = 1). Polarised scores may increase the strength of correlations between scores  Although statistically significant a change in pain score from 0 to 1 of questionable clinical significance  FLACC correlated better with CAS scores (pain assessment) than FAS scores (distress assessment) suggesting FLACC assesses pain better than distress in this age group.  Correlation between FLACC and self-report slightly higher in younger children. May suggest that older children demonstrate different or suppressed pain behaviours |
| Ranger et al, 2013 [[17](#_ENREF_17)] | To determine whether noxious stimuli is associated with regional cerebral hemodynamic changes and whether these changes correlate with physiologic and behavioural measures in critically ill infants with congenital heart defects during chest-drain removal after cardiac surgery.  Repeated measures study  NIRS monitoring during painful procedure which was also video-taped for later assessment by pain nurse specialist using the FLACC scale. NIRS and FLACC scores analysed for 3x 30sec epochs addressing:   * Baseline * Tactile stimulation * Removal of drain   Assessments repeated 10 weeks later and a second assessor applied FLACC scores to half of the videos | 20 critically unwell infants less than 12 months  Setting: Cardiac ICU,  Pain: postoperative, procedural  Index: NIRS, FLACC  Reference: na | Inter-rater: For FLACC scores ICC = 0.86 (0.71 – 0.960)  Intra-rater: For FLACC scores ICC = 0.9.  COSMIN - poor | Hypothesis (convergent): No association between FLACC scores or its 5 items and NIRS or physiological signal changes across the different time periods.  Significant change in MAP (F1.38, 16.52 – 19.18, p,0.001) and HR (F1.28, 24.27 = 6.87, p=0.01) across epochs  COSMIN – poor  Responsiveness: Mean FLACC pain scores at the 3 epochs were: baseline 0.25 (SD 0.12), 95% CI [0.01, 0.51]; tactile 3.25 (SD 0.56), 95% CI [2.08, 4.23]; and noxious 6.7 (SD 0.66), 95% CI [5.32, 8.08]. Overall, FLACC scores differed significantly between epochs (F1.19=102.64; p<0.001) ANOVA  Analgesic administration associated with reduced change in HbO2 (p=0.005), HbH (p=0.002) and HR (p=0.02) in response to noxious stimuli but no significant impact on FLACC score. Sedation associated with less change in HbO2 (p = 0.017), bilateral HbH (p = 0.015 & 0.008), HR (p = 0.005) and SpO2 (p = 0.012) and FLACC score (p = 0.005)  COSMIN - poor | Not assessed | Increase in FLACC scores and across epochs blunted by administration of sedation but not analgesic (high FLACC scores during despite analgesic) and no association between FLACC and NIRS - suggests FLACC may be measuring constructs in addition to or instead of pain/ insufficiently sensitive to detect pain in critically unwell neonates.  Potential that all scale items can’t be adequately assessed from video footage. |
| Suraseranivongse et al. 2002 [[18](#_ENREF_18)] | To assess: 1) agreement & correlation of nurses & parent scores, 2) difference between CHEOPS & FLACC scores as rated by nurses and parents, 3) effects of parental education on scores and 4) ease with which parents could score CHEOPS & FLACC.  Descriptive comparison study  Parents trained to use CHEOPS and FLACC  One of 2 nurses [tested for inter-rater reliability (ICC > 0.9)] and parent (blinded) scored pain of child in the recovery room at 15minutely intervals using CHEOPS and FLACC | Parents of 69 children, age 1-12 years (median 3.7, IQR 2-6.85 years)  Setting: Recovery room  Pain: Post-operative (herniorrhaphy or hydrocelectomy)  Index: parent FLACC and CHEOPS (Thai)  Reference: nurse FLACC & CHEOPS (Thai) | Inter-rater: High correlation between nurses & parents using FLACC (rho = 0.938, p < 0.001) & CHEOPS (rho = 0.945, p < 0.001) Agreement ICC = 0.949 & 0.977 respectively (p < 0.001).  No difference between parent scores and nurse scores for FLACC (p = 0.166) or CHEOPS (0.544)  COSMIN - fair | Not assessed | Feasibility: Ease of use scores (VAS) for FLACC (3.38 ± 1.70) and CHEOPS (3.43 ± 1.75) were not different (p = 0.815).  \* Authors claim ease of use demonstrated by reliable application by parents (regardless of level of education) | Sequencing of application of scales not defined but applied at the same time - hence may impact on scores for scales.  Pain scores very low (median FLACC = 0, IQ = 0 - 3). Polarised scores may increase the strength of correlations between scores |
| [(Suraseranivongse, Santawat et al. 2001)](#RANGE!_ENREF_90) | To: 1. cross-validate scales (FLACC, CHEOPS, OPS and TPPPS in Thai children, measuring validity, reliability and practicality and 2. assess the discriminative ability of the scales  Cross validation study  Scale translated and tested for content validity by experts  Behaviours before &after surgery, before & after analgesics videotaped, sequence randomly arranged then 4 observers scored using 4 scales  2 weeks later observers rescored 30 behaviours | 167 children aged 1 – 5.5yr  Setting: PACU and surgical ward  2 centres, Thailand  Pain: Post-operative  Index: FLACC, CHEOPS, OPS and TPPPS (Thai)  Reference: na | Inter-rater: CHEOPS = 0.9184, OPS = 0.9198  TPPPS = 0.9657, FLACC = 0.9488  Intra-rater: CHEOPS = 0.99 – 0.87, OPS = 0.95 – 0.99, TPPPS = 0.92 – 0.99, FLACC = 0.95 – 0.99  COSMIN - good | Content (face): FLACC accepted unchanged. 2 behaviours in CHEOPS and 1 in TPPPS opposed as not seen in Thai children.  COSMIN - poor  Hypothesis (convergent): correlations btw scales 0.62 – 0.77 (p < 0.0001) CHEOPS/OPS highest, FLACC/TPPPS lowest.  COSMIN - fair  Responsiveness: significant change in scores before and after surgery before analgesic for all scales p < 0.001.  COSMIN - good | Feasibility - Time taken to apply scales FLACC – 45.5s, CHEOPS – 59s, OPS – 44s & TPPPS- 40.1s,.  Scales ranked by 76.7 – 90% (FLACC – 90%) as ‘feasible for clinical use’, 73.3% (FLACC, OPS & TPPPS) & 80% (CHEOPS) as ‘easy to use’, 80 – 100% (FLACC = 86.7%) as helpful for assessment & 26.6 – 66.7 % (FLACC = 66.7%) as generally satisfied  Utility: cut off & decision to treat pain: CHEOPS highest cut off = 6 and strongest agreement with intention to Rx by clinician, к = 0.83, FLACC cutoff = 2, к = 0.659). | Clinical intent to treat used to establish scale cut offs – pain diagnosed by nurse but analgesic decision made by researcher ‘blinded to score’ but aware that nurse diagnosed ‘pain’. Those not diagnoses by nurse not referred to researcher for decision to treat therefore biasing cohort for treatment  Potential that all scale items can’t be adequately assessed from video footage.  Scoring of multiple tools may impact on convergence |
| Taddio et al, 2011 [[19](#_ENREF_19)] | To investigate the reliability, validity and practicality of 3 observational measures of acute pain for the assessment of pain in infants undergoing vaccine injections.  Convenience sample from an RCT.  Infants having 1st vaccination in clinical trial comparing pain associated with two vaccines (DPTaP-Hib or PCV). Videotaped & pain scored at baseline and 15sec after vaccination from video  Phase 1: single raters scored all infants using all 3 scales. 2nd rater scored 30 randomly selected infants  Phase 2: 3 different raters applied scale after one view of video. Scored again after watching video as often as required to score confidently.  All raters surveyed about utility of scales | 120 infants aged 2 – 6mth  Setting: private outpatient practice, Canada  Pain: Immunisation  Index: FLACC, MBPS and NIPS  Reference: na | Inter-rater: ICC > 0.85 for pre and post vaccination for all scales, FLACC – 0.85 and 0.94, NIPS – 0.9 and 0.92, and MBPS – 0.94 and 0.9 respectively.  Intra-rater: FLACC (ICC, 0.98: 95% CI, 0.97–0.99), MBPS (ICC, 0.96: 95% CI, 0.94–0.97) & NIPS (ICC, 0.98:95%CI, 0.97–0.98)  COSMIN - good  Internal consistency: cronbach’s alpha > 0.83 for all scales at baseline & following vaccination  COSMIN - poor | Criterion (concurrent)\*: Pearson correlation btw scales MBPS & FLACC (r=0.84), FLACC & NIPS (r=0.92) (p<0.001) &MBPS & NIPS (r=0.87) (p<0.001)  COSMIN - poor  Hypothesis (discriminant): Scores lower for all scales (p<0.001) for infants receiving DPTaP-Hib to infants receiving PCV. FLACC (5.3 versus. 7.8, p< 0.001), MBPS (6.8 versus 8.5, p<0.001) & NIPS (4.4 versus 6.2, p<0.001)  COSMIN - fair  Responsiveness: significant increase scores for all scales pre and post vaccination (p<0.001), FLACC (0.6 versus 6.5), MBPS (2.3 versus 7.7) & NIPS (0.3 versus 5.3).  COSMIN - good | Feasibility: Agreement (ICC) for first score and final score, high for: FLACC (ICC, 0.98: 95% CI, 0.97–0.99), MBPS (ICC, 0.96: 95% CI, 0.94–0.97) & NIPS (ICC, 0.98:95%CI, 0.97–0.98)  Percentage of pain assessments recorded after one viewing did not differ significantly (p=0.06) among groups: MBPS (56%), NIPS (66%), FLACC (50%).  Total time taken to assess pain lowest for MBPS (5h 25min), followed by the NIPS (5h 58min) and FLACC (6h 50min).  User preference highest for MBPS (80%) | Correlations reported as criterionvalidity - observational scale used as reference – not a gold standard  Potential that all scale items can’t be adequately assessed from video footage. |
| Voepel-Lewis et al, 2010 [[23](#_ENREF_23)] | To evaluate the reliability and validity of the Face, Legs, Activity, Cry, Consolability (FLACC) Behavioural Scale in assessing pain in critically ill adults and children unable to self-report pain.  Observational design.  3 nurses independently, observed and scored pain behaviours (2 using FLACC and 1 Checklist of Nonverbal Pain  Indicators (for adults) or the COMFORT scale (for children) before analgesic or during a painful procedure, and 15 to 30 minutes after analgesic or the procedure concluded | 37 critically ill patients (8 children aged 5.6 years) to (results pooled for adults and children)  Setting: ICU and PICU  Pain: acute pain pre-analgesic or procedural  Index: FLACC  Reference: COMFORT scale, NVPI | Inter-rater: (children) - exact agreement (58% - 83%), κ statistics (0.33 – 0.71) and ICC (0.43 -0.92) for items and total FLACC = 0.85 (CI: 0.52 – 0.96)  COSMIN - Good  Internal consistency: Cronbach α = 0.882 when all items included. Improved (00.934) with removal of cry  COSMIN - fair | Criterion (concurrent): FLACC scores correlated highly with NVPI and COMFORT scores (rho = 0.963 and 0.849, respectively)  COSMIN - poor  Responsiveness: decreases in FLACC scores after analgesic (or from painful to non-painful) (mean=5.27; SD, 2.3 versus mean= 0.52; SD, 1.1; p<.001)  COSMIN - fair | Not assessed | Correlations reported as criterion validity - observational scale used as reference – not a gold standard  Author a member of original scale development and validation study team |
| Voepel\_Lewis et al, 2008 [[21](#_ENREF_21)] | To evaluate pragmatic attributes or clinical utility properties of 3 recently developed pain assessment tools for children with CI  Observational  Clinicians scored 15 videotaped observations recorded during previous study over first 3 postoperative days  Applied scale to 5 video segments each.  Completion of Clinical Utility Attributes Questionnaire (CUAQ) | 20 clinicians (5 physicians and 15 nurses)  15 cognitively impaired children  Setting: not documented  Pain: Post-operative  Index: r-FLACC (modified descriptors), NAPI & NCCPC-PV  Reference: na | Interrater: ICC between participant scores and originally assigned scores range from -0.06 (NCCPC-PV to NAPI) to 0.92 (r-FLACC to r-FLACC) and kappa scores > 0.71 for all scores (r-FLACC to r-FLACC highest at 0.96)  COSMIN - poor | Not assessed | Utility: CUAQ scores higher for r-FLACC (49.6 ± 4.6) and NAPI (43.7 ± 6.7) compared to NCCPC-PV (24.9 ± 8.1.)  Feasibility: time taken to score the r-FLACC and NAPI (2.9 ± 1.7 and 2.8 ± 1.5) shorter than for NCCPC-PV (5.1 ± 2.2 min) p < 0.001) | Participants familiar with several assessment tools but unfamiliar with NAPI and NCCPC-PV which may influence assessment of ease in application.  Potential that all scale items can’t be adequately assessed from video footage.  Author a member of original scale development and validation study team |
| Voepel\_Lewis et al, 2005 [[20](#_ENREF_20)] | To evaluate the validity of parental pain scores of children with CI  Observational design  Parents individualised FLACC descriptors before application  Post-operatively children scored independently by 2 nurses and parent (who also scored using NRS) using FLACC  Observations repeated following analgesics for those who received analgesics | 52 children aged 4 – 19years (mean = 11.3 ±4.7 years)  with cognitive impairment  Setting: not documented  Pain: Post-operative  Index: parent applied FLACCr (modified descriptors)  Reference: simplified FACES Scale, observer NRS | Not reported (2 nurses applied FLACC independently but data not reported)  Agreement between parents and nurses FLACC scores (ICC = 0.78; CI = 0.63–0.87) and parents’ NRS rating and nurse FLACC scores (ICC = 0.73; CI = 0.59–0.83). Parent scores tended to be slightly higher than the nurse ratings (FLACC bias = 0.98± 2.2; NRS bias = 1.45 ± 2.2).  COSMIN - good | Criterion (concurrent): Agreement between child’s rating and parents’ FLACC score (к=0.43),  COSMIN - poor  Hypothesis (convergent): correlation between parents’ NRS score and FLACC scores - (ICC = 0.81; CI = 0.70–0.89). NRS scores tended to be higher than their FLACC ratings (bias = 0.56 ± 1.82), nurses FLACC score (к=0.65) & parents NRS score (к=0.39)  COSMIN - fair  Responsiveness: Parental assessment of pain decreased after analgesics (FLACC 6.6 ± 2.4 vs. 2.6±2; p = .003; NRS: 6.4 ± 2.5 vs. 3.1 ± 2.3; p = .004)  COSMIN - fair | Not assessed | Small numbers of children able to self-report n = 12  FLACC reaches acceptable levels but lower than nurses or parents NRS  Author a member of original scale development and validation study team |
| Voepel-Lewis et al, 2002 [[22](#_ENREF_22)] | To evaluate the validity and reliability of the Face, Legs, Activity, Cry, Consolability (FLACC) tool for assessing pain in children with CI  Observational design  The child’s nurse observed and scored pain with the FLACC tool before &after analgesics. Parents also scored pain using a VASobs, and where possible children self-reported pain  Observations videotaped for FLACC scoring by 2 nurses blinded to analgesics & real time pain scores. Reassessed 50 randomly selected observations 2-3 months later.  Pain scores were coded as mild (0–3), moderate (4–6), and severe (7–10). | 79 cognitively impaired children aged 4 – 18 (mean 10.11 ± 4.3 yr) resulting in 140 observations.  Setting: Bedside  Pain: Post-operative (orthopaedic or general surgery)  Index: FLACC  Reference: VAS obs, simplified FACES or NRS for self-report | Inter-rater: Moderate agreement between all observers for total scores (r = 0.507-0.778 p≤0.0001) and for each FLACC item (0.339 – 0826, p≤0.0001). Measures of exact agreement highest between blinded observers for all categories (к scores, face = 0.346, legs = 0.477, activity = 0.405, cry = 0.652, consolability = 0.555)  Intra-rater: correlations for test-retest FLACC scores for the blinded observers (r = 0.8–0.883; p< 0.001)  COSMIN - fair | Hypothesis (convergent): Correlation between FLACC scores of bedside nurse (r113, 0.651) & both blinded nurses (r94 = 0.609 and 0.519) with parent scores significant (p< 0.001),  Parent scores higher than FLACC scores of bedside nurse (bias 0.59, precision ±2.3) and blinded nurses (0.51 ±2.4 and 0.65 ±2.6). Bedside nurses scores higher than blinded nurses (0.2 ±1.6 and 0.09 ±2.4)  COSMIN - poor  Responsiveness: Decrease in FLACC scores after analgesics (5.3 ± 2.8 versus 2.0 ± 2.4 for the bedside nurses’ scores, p < 0.001; 5.1 ± 2.9 versus 2.2 ± 3.0 for the blinded nurses’ scores, p = 0.001)  COSMIN - fair | Utility: excellent greement for mild and severe pain categories (and good agreement for moderate pain. Children with mild pain most often received no analgesic (64%) or non-opioids (18%), those with moderate to severe pain most often received morphine (60%) or diazepam (6%) for muscle spasms | Potential that all scale items can’t be adequately assessed from video footage.  Video observers blinded – reducing the bias supports the influence of clinical information as scores consistenyly lower than both bedside nurses and parents scores  Author a member of original scale development and validation study team |

| **Study** | **Study aim / Study design** | **Sample/Setting/ Pain experience** | **Reliability** | **Validity** | **Feasibility and clinical utility** | **Comments** |
| --- | --- | --- | --- | --- | --- | --- |
| Alternate index scale (FLACC as reference scale - concurrent validity testing) [5] | | | | | | |
| Chorney et al, 2011 [[4](#_ENREF_4)] | To develop and validate a behavioral coding measure, the Children’s Behavior Coding System-PACU (CBCS-P), for children’s distress and nondistress behaviors while in the post-anesthesia recovery unit.  Observational study  Children videoed and other scales reviewed to identify scale items  Nurse pain scores (using FLACC) and analgesic use recorded from medical record.  Children videoed postoperatively, trained observers scored videos using CBCS-P. | 121 children aged 2 - 11 years and their parents  Setting: PACU  Pain: Postoperative  Index: CBCS-P | FLACC not assessed | Hypothesis (convergent) Verbal Distress Composite score was significantly positively correlated with FLACC (r=0.24, p<.05), and analgesic use (r=0.28, p<.01), as was Nonverbal Distress Composite score (FLACC, r=0.40, p<.001, analgesic, r=0.30, p<.001).  COSMIN - fair | FLACC not assessed | FLACC used as reference to test convergent validity of new unvalidated scale – therefore can not provide evidence of FLACC validity |
| Fournier-Charriere et al, 2012 [[7](#_ENREF_7)] | To elaborate and validate a tool, relevant in any painful situation, with agitation or prostration, and for any age under 7 years  Observational study  Parent, clinical nurse and researcher assessed pain (global score), need for analgesic and the level of fatigue & anxiety  Pain scored using EVENDOL(CHEOPS, TPPPS, FLACC and EDIN) by nurse and researcher and child>4years at:  T1R – triage at rest  T1M – triage during examination  Post analgesic  T2R – at rest  T2M – during examination  Subset of children filmed for reliability assessment using 6 assessors | 291 children aged 0 – 83 months  37 video recordings  Setting: Emergency department (4 centres in France)  Pain: Procedural and acute pain  Index: EVENDOL  Reference: FLACC, CHEOPS, EDIN & TPPPS | FLACC not assessed | Hypothesis (convergent): correlations with other scales (incl. FLACC) all >0.7 at T1 (FLACC not reported separately)  COSMIN - poor | FLACC not assessed | FLACC used as reference to test convergent validity of new unvalidated scale – therefore can not provide evidence of FLACC validity  FLACC results not described separately from other BPS therefore doesn’t contribute to FLACC validation  Scoring of multiple tools may impact on convergence |
| Hatrrick et al, 2002 [[9](#_ENREF_9)] | To evaluate the reliability of the TPPPS following general anaesthesia as a guide to analgesic administration, and to compare the validity of this scale with other observational measures of pain in children  Prospective, consecutive, single-blinded observation study.  Two researchers blinded to each others’ assessments applied scale(s).  All scales applied by all assessors to subset of sample to detect changes in pain following analgesic | 51 children\* assessed aged 1.1 – 5.1 (mean 2.9 years)  Setting: operative & radiology suites  Pain: Post-operative & non-painful radiological procedures  \*12 children scored using all scales to detect changes pre and post analgesics  Index: TPPPS  Reference: FLACC, modified COMFORT\* | FLACC not assessed | Responsiveness: Significant decreases in scores detected following analgesic for all scales. (FLACC p<.002) No decrease for children who did not receive analgesic.  COSMIN - poor | FLACC not assessed | Observers not blinded to circumstances |
| Ramlet et al, 2007 [[15](#_ENREF_15)] | To test the preliminary psychometric properties of the Multidimensional Assessment Pain Scale (MAPS), a tool for assessing postoperative pain in critically ill preverbal children.  Repeated measures design  Tool developed by authors and validated using FLACC and VASobs  Up to 8 measurements of pain using FLACC, MAPS and VASobs were made by 2 independent clinical nurses over a maximum period of 48 h post-surgery. | 43 critically unwell children aged 0 – 31 months (median = 9.5)  Setting: PICU  (Australia & Switzerland)  Pain: post-operative  Index: MAPS  Reference: FLACC | FLACC not assessed | Hypothesis (convergent): Bland and Altman method to determine the agreement between MAPS & FLACC scales and MAPS & VASobs. The mean of the differences were 0.44 (CI: 0.18–0.71) and 0.25 (CI: .0.02 – 0.49) respectively. The limits of agreement were -1.22 to 2.09 and -1.24 to 1.74.  COSMIN - poor | Not applicable | FLACC used as reference to test convergent validity of new unvalidated scale – therefore can not provide evidence of FLACC validity |
| Ramlet et al 2007 [[16](#_ENREF_16)] | To evaluate the clinical validity and utility of the MAPS.  Repeat measures study  Pain measured by 2 observers using FLACC and MAPS before and after analgesic. Clinicians independently made treatment decision and assessed pain using VASobs | 20 critically ill children aged between 4 days and 31 months (median=7.5 months)  Setting: PICU  Pain: post-operative  Index: MAPS  Reference: FLACC | FLACC not assessed | Convergent: Bland and Altman method to determine the agreement between MAPS and FLACC scales and MAPS and VASobs. The mean of the differences were - 0.12 and -0.29 respectively. The limits of agreement were - 3.71 to 3.78 and 1.78 to -2.37.  COSMIN - poor | FLACC not assessed | FLACC used as reference to test convergent validity of new unvalidated scale – therefore can not provide evidence of FLACC validity |

### Note

CAS - Colour Analogue Scale, CBCS-P - Children’s Behavior Coding System-PACU, CHEOPS – Children’s Hospital of Eastern Ontario Pain Scale, CHIPPS - Children’s and Infants’ Postoperative Pain Scale, CICU – cardiac intensive care unit, CNPI – Checklist of Non Verbal Pain Indicators, CPOT - Critical Care Pain Observation Tool, CRIES – Cry, Requires O2, Increased vital signs, Expression, Sleeplessness, FAS - Facial Affective Scale, FASS – Facial Action Summary Score, MAPS - Multidimensional Assessment Pain Scale, MBPS - Modified Behavioural Pain Scale, NAPI – Nursing Assessment of Pain Intensity, NIPS – Neonatal Pain Indicators, NCCPC-PV - Non-communicating Children’s Pain Checklist Postoperative Version, NVPS – (adult) Non-verbal Pain Scale, OBS – Objective Pain Scale, PAINAID – Pain Assessment in Advanced Dementia, PIPP – Premature Infant Pain Profile, TPPPS - Toddler, Pre-schooler Postoperative Pain Scale, **VAS – Visual Analogue Scale**

## References

[1] Ahn Y, Jun Y. Measurement of pain-like response to various NICU stimulants for high-risk infants. Early Human Development 2007;83(4):255-262.

[2] Bai J, Hsu L, Tang Y, van Dijk M. Validation of the COMFORT Behavior scale and the FLACC scale for pain assessment in Chinese children after cardiac surgery. Pain Management Nursing 2012;13(1):18-26.

[3] Bringuier S, Picot M-C, Dadure C, Rochette A, Raux O, Boulhais M, Capdevila X. A prospective comparison of post-surgical behavioral pain scales in preschoolers highlighting the risk of false evaluations. Pain 2009;145(1-2):60-68.

[4] Chorney JM, Tan ET, Martin SR, Fortier MA, Kain ZN. Children's Behavior in the Postanesthesia Care Unit: The Development of the Child Behavior Coding System-PACU (CBCS-P). Journal of Pediatric Psychology 2011.

[5] da Silva FC, Santos Thuler LC, de Leon-Casasola OA. Validity and reliability of two pain assessment tools in Brazilian children and adolescents. Journal of Clinical Nursing 2011;20(13/14):1842-1848.

[6] da Silva FC, Thuler LCS. Cross-cultural adaptation and translation of two pain assessment tools in children and adolescents. Jornal de Pediatria 2008;84(4):344-349.

[7] Fournier-Charrière E, Tourniaire B, Carbajal R, Cimerman P, Lassauge F, Ricard C, Reiter F, Turquin P, Lombart B, Letierce A, Falissard B. EVENDOL, a new behavioral pain scale for children ages 0 to 7&#xa0;years in the emergency department: Design and validation. Pain 2012;153(8):1573-1582.

[8] Gomez RJ, Barrowman N, Elia S, Manias E, Royle J, Harrison D. Establishing intra- and inter-rater agreement of the face, legs, activity, cry, consolability scale for evaluating pain in toddlers during immunization. Pain Research and Management 2013;18(6):e124-e128.

[9] Hartrick CT, Kovan JP. Pain assessment following general anesthesia using the Toddler Preschooler Postoperative Pain Scale: a comparative study. Journal of Clinical Anesthesia 2002;14(6):411-415.

[10] Johansson M, Kokinsky E. The COMFORT behavioural scale and the modified FLACC scale in paediatric intensive care. Nursing in Critical Care 2009;14(3):122-130.

[11] Malviya S, Voepel-Lewis T, Burke C, Merkel S, Tait AR. The revised FLACC observational pain tool: improved reliability and validity for pain assessment in children with cognitive impairment. Paediatric Anaesthesia 2006;16(3):258-265.

[12] Manworren RCB, Hynan LS. Clinical validation of FLACC: preverbal patient pain scale. Pediatric Nursing 2003;29(2):140-146.

[13] Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. Pediatric Nursing 1997;23(3):293-297.

[14] Nilsson S, Finnstrom B, Kokinsky E. The FLACC behavioral scale for procedural pain assessment in children aged 5-16 years. Paediatric Anaesthesia 2008;18(8):767-774.

[15] Ramelet A-S, Rees N, McDonald S, Bulsara M, Abu-Saad HH. Development and preliminary psychometric testing of the Multidimensional Assessment of Pain Scale: MAPS. Paediatric Anaesthesia 2007;17(4):333-340.

[16] Ramelet A-S, Rees NW, McDonald S, Bulsara MK, Huijer Abu-Saad H. Clinical validation of the Multidimensional Assessment of Pain Scale. Paediatric Anaesthesia 2007;17(12):1156-1165.

[17] Ranger M, Celeste Johnston C, Rennick JE, Limperopoulos C, Heldt T, du Plessis AJ. A Multidimensional Approach to Pain Assessment in Critically Ill Infants During a Painful Procedure. Clin J Pain 2013.

[18] Suraseranivongse S, Kraiprasit K, Petcharatana S. Postoperative pain assessment in ambulatory pediatric patients by parents. Journal of the Medical Association of Thailand 2002;85 Suppl 3:S917-922.

[19] Taddio A, Hogan M-E, Moyer P, Girgis A, Gerges S, Wang L, Ipp M. Evaluation of the reliability, validity and practicality of 3 measures of acute pain in infants undergoing immunization injections. Vaccine 2011;29(7):1390-1394.

[20] Voepel-Lewis T, Malviya S, Tait AR. Validity of parent ratings as proxy measures of pain in children with cognitive impairment. Pain Management Nursing 2005;6(4):168-174.

[21] Voepel-Lewis T, Malviya S, Tait AR, Merkel S, Foster R, Krane EJ, Davis PJ. A comparison of the clinical utility of pain assessment tools for children with cognitive impairment. Anesthesia & Analgesia 2008;106(1):72-78, table of contents.

[22] 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. Anesthesia & Analgesia 2002;95(5):1224-1229, table of contents.

[23] Voepel-Lewis T, Zanotti J, Dammeyer JA, Merkel S. Reliability and validity of the face, legs, activity, cry, consolability behavioral tool in assessing acute pain in critically ill patients. American Journal of Critical Care 2010;19(1):55-61; quiz 62.

[24] Willis MHW, Merkel SI, Voepel-Lewis T, Malviya S. FLACC Behavioral Pain Assessment Scale: a comparison with the child's self-report. Pediatric Nursing 2003;29(3):195-198.