



ความเท่าเทียมกันของการสังเกตและความเป็นไปได้ในการนำไปใช้เครื่องมือประเมินความปวด
คอมฟอร์ทบีและแฟลค ฉบับปรับปรุงภาษาไทย หลังได้รับความรู้และการฝึกทักษะ
ของพยาบาลในเด็กไทยหลังผ่าตัดที่ใส่ท่อช่วยหายใจ¹

INTER-RATER RELIABILITY AND FEASIBILITY OF USING THE COMFORT-B (Thai)
AND THE MODIFIED FLACC (Thai) PAIN SCALE BY NURSES AFTER EDUCATION
AND SKILL TRAINING IN POST-SURGICAL PAIN FOR VENTILATED THAI CHILDREN

วันฉวี วิรุฬห์พานิช²
Wantanee Wiroonpanich

กาญจนา นาคทิม³
Kanchana Narktim

ปิยธิดา เทพประดิษฐ์⁴
Piyatida Theppradit

¹วิจัยได้รับทุนจากคณะพยาบาลศาสตร์ มหาวิทยาลัยสงขลานครินทร์

This research was funded by the Faculty of Nursing, Prince of Songkla University.

²ผู้ช่วยศาสตราจารย์ ดร. คณะพยาบาลศาสตร์ มหาวิทยาลัยสงขลานครินทร์ สงขลา 90110

Assistant Professor, Ph.D., Faculty of Nursing, Prince of Songkla University, Songkhla, 90110, Thailand

³พยาบาลวิชาชีพ โรงพยาบาลหาดใหญ่ หาดใหญ่ สงขลา 90110

Registered Nurse, Hatyai Hospital, Hatyai, Songkhla, 90110, Thailand

⁴อาจารย์พยาบาล คณะพยาบาลศาสตร์ มหาวิทยาลัยสงขลานครินทร์ สงขลา 90110

Nursing Instructor, Faculty of Nursing, Prince of Songkla University, Songkhla, 90110, Thailand

Corresponding author E-mail: wantanee.w@psu.ac.th

Received: September 8, 2020

Revised: March 13, 2021

Accepted: March 19, 2021

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อประเมินความเท่าเทียมกันของการสังเกตการใช้เครื่องมือประเมินความปวด
คอมฟอร์ทบีและแฟลค ฉบับปรับปรุงภาษาไทย และความเป็นไปได้ในการใช้เครื่องมือประเมินความปวดทั้งสองชนิด
หลังจากได้รับความรู้และการฝึกทักษะของพยาบาลในการประเมินความปวดในเด็กหลังผ่าตัดใส่ท่อช่วยหายใจ
ตัวอย่าง คือ ผู้ทรงคุณวุฒิกับพยาบาลจำนวน 19 คน ในหอผู้ป่วยวิกฤตเด็ก โดยพยาบาลจะได้รับการอบรมความรู้
เกี่ยวกับความปวด เครื่องมือประเมินความปวด และฝึกทักษะการประเมินความปวดโดยใช้เครื่องมือประเมินความ
ปวดทั้งสองชนิดในเด็กหลังผ่าตัดที่ใส่ท่อช่วยหายใจจากวิดีโอและสถานการณ์จริงในหอผู้ป่วย จากนั้นพยาบาลแต่
ละคนทำการสังเกตและประเมินความปวดเด็กหลังผ่าตัดและใส่ท่อช่วยหายใจจากชุดวิดีโอที่ผู้วิจัยเตรียมไว้ พยาบาล
แต่ละคนมีอิสระในการสังเกตแต่ต้องประเมินความปวดในเวลาเดียวกัน โดยใช้เครื่องมือประเมินความปวดทั้งสองชนิด
ข้อมูลที่ได้นำมาวิเคราะห์ความเที่ยงของความเท่าเทียมกันของการสังเกตระหว่างผู้ประเมินและตัวอย่างด้วยสถิติ
แคปปาของโคเฮน และความเป็นไปได้ในการใช้เครื่องทั้งสองชนิดโดยพิจารณาจากค่าร้อยละ

ผลการวิจัยพบว่า คะแนนสัมประสิทธิ์แคปปาโดยรวมของเครื่องมือประเมินความปวดคอมฟอร์ทบีและแฟลคฉบับปรับปรุงภาษาไทยระหว่างผู้ทรงคุณวุฒิและพยาบาลเท่ากับ .86 และ .89 ตามลำดับ ซึ่งคะแนนที่ได้จัดอยู่ในระดับดี สำหรับความเป็นไปได้ในการนำไปใช้ของเครื่องมือประเมินความปวดแฟลคฉบับปรับปรุงภาษาไทย พบว่าพยาบาลให้คะแนน 3 และ 4 รวมกันร้อยละ 100 ทุกข้อคำถาม และความเป็นไปได้ในการนำไปใช้ของเครื่องมือประเมินความปวดคอมฟอร์ทบี ฉบับปรับปรุงภาษาไทย พบว่าพยาบาลให้คะแนน 3 และ 4 ต่ำกว่าร้อยละ 70 ในข้อคำถามการนำไปใช้ในการประเมินความปวดในเด็กหลังผ่าตัดและใส่เครื่องช่วยหายใจได้เร็วและง่าย ดังนั้น ผลการศึกษาครั้งนี้ชี้ให้เห็นว่าการอบรมเชิงปฏิบัติการด้วยการสอนให้ความรู้และการฝึกทักษะการใช้เครื่องมือประเมินความปวดทั้งสองชนิดจนกระทั่งพยาบาลได้คะแนนความเท่าเทียมกันของการสังเกตผ่านเกณฑ์เป็นที่ยอมรับ คือ มากกว่าหรือเท่ากับร้อยละ 80 จะช่วยให้พยาบาลเพิ่มพูนความรู้และทักษะในการใช้เครื่องมือประเมินความปวดในการปฏิบัติการพยาบาลในคลินิกได้อย่างเพียงพอและเหมาะสม

คำสำคัญ: ความเท่าเทียมกันของการสังเกต, ความเป็นไปได้ในการนำไปใช้, คอมฟอร์ทบี, แฟลค

Abstract

This study aimed to evaluate the inter-rater reliability and feasibility of the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale by nurses after education and skill training in assessing post-surgical pain in ventilated Thai children between a standardized person and nurses in a Pediatric Intensive Care Unit (PICU). Nineteen nurses received on-ward training from the researcher in pediatric pain knowledge and skills and how to use both scales in assessing post-surgical pain in ventilated children from videos and the real situation. Then, each nurse simultaneously and independently observed and evaluated post-surgical pain in ventilated children from sets of videos using these scales. Interrater reliability between the standardized person and nurses was analyzed by Cohen's Kappa, and the feasibility of using these two scales was analyzed by percentage.

Results revealed that the overall Kappa scores of the COMFORT-B (Thai) and the modified FLACC (Thai) scale between the standardized person and all nurses were .86 and .89, respectively, which can be categorized as very good. For the feasibility of the modified FLACC (Thai) scale, all items were rated as either 3 or 4 by 100% of nurses. Nurses' agreement with the feasibility of the COMFORT-B (Thai) was rated as either 3 or 4 by < 70% in the items of the scale being quick and easy to use to assess post-surgical pain in ventilated children. These findings indicate that teaching and workshop training on how to use these scales until nurses meet the acceptable agreement percentage of $\geq 80\%$ will enhance their knowledge and skills to use the pain scales adequately in clinical practice.

Keywords: inter-rater reliability, feasibility, The COMFORT-B (Thai), The modified FLACC (Thai)

Background and significance

Pain in children is hard to assess, especially post-surgical pain in intubated and ventilated children in the pediatric intensive care unit (PICU) because these

children are unable to self-report the level of pain they experienced. Some children can express pain by showing behavior responses (such as facial expression, crying, or body movement), and physiological

changes (such as heart rate, respiratory rate, or oxygen saturation). However, some may not be able to express their pain through behavior due to cognitive development, the severity of illness, or the difficulty in differentiating between pain and other constructs such as stress, agitation, and sedation (Ambuel, Hamlett, Marx, & Blumer, 1992; Ramelet, Abu-Saad, Rees, & McDonald, 2004). Thus, the assessment of pain in ventilated children depends on observation and evaluation by health care providers (Puntillo, Miaskowski, Kehrl, Gleeson, & Nye, 1997). If health care providers do not clearly understand these behaviors expressed by ventilated children, under- or over-assessment of the pain level may lead to poor pain management and suboptimal treatment (Manworren & Stinson, 2016). The consequences of untreated pain include an increase in catecholamine and stress hormone levels, which are a potential cause of physiological and psychological distresses such as tachycardia, hypertension, increased oxygen requirement, and decreased tissue perfusion (Ambuel et al., 1992; Blakely & Page, 2001; Hamill-Ruth & Marohn, 1999; Kwekkeboom & Herr, 2001). As well as a severe clinical complication, delayed recovery increases the risk of mortality and long-term psychological consequences such as anxiety and fear (Linton & Shaw, 2011). One factor that contributes to poor pain management in this children group is the difficulty in assessing pain (Ismail, 2016). Thus, accurate pain assessment is a key component to estimate ventilated children's pain, which in turn leads to appropriate pain management.

To assess pain as accurately as the experience of ventilated children, pain scales using behavioral and physiological indicators have been developed. The most common behavioral pain scales used in the PICU are the COMFORT-B scale and the modified FLACC scale (Oakes, 2011). These instruments have been shown to have good validity and reliability in assessing the pain of ill children, particularly in post-surgical ventilated children in critical care settings (Bai, Hsu, Tang, & Dijk, 2012; Johansson & Kokinsky, 2009; van Dijk et al., 2000; Voepel-Lewis, Merkel, Tait, Trzcinka, & Malviya, 2002). These scales are categorized as behavioral categories and items with the definition of each item for scoring pain. Each behavioral item is scored on a Likert scale depending upon different pain behaviors. Although each item is simply defined, it is difficult for health care providers to evaluate the quality of each behavior if they do not understand children's behaviors because of pain and factors influencing them in a ventilated environment. Therefore, education specifically on behavioral pain assessment of ventilated children with critical illness is needed, and it has been suggested by the expert that improving nursing pain management requires more than knowledge acquisition (Watt-Watson, 1997).

Presently, nursing education programs for pain management have been developed and the outcomes have indicated improvements in nurses' pain knowledge, beliefs, attitudes, and management of clinical skills specifically in pain assessment

in a hospital environment (Hurley-Wallace, Wood, Franck, Howard, & Lioffi, 2019; Michaels, Hubbartt, Carroll, & Hudson-Barr, 2007). Examples of teaching methods are interactive pain lectures, on-ward skill development to enhance pain scale use, assessment guide, web-based instruction, and feedback. In Thailand, pain management is included in the nursing teaching curriculum. However, the time allocated for teaching is only about 2-3 hours and the content is focused mostly on pain theory, mechanism, pain scales, and pharmacological and non-pharmacological pain management without any training on how to use the pain scales or non-pharmacological pain management. Thus, basic pain education for nursing students is not enough for effective pain assessment and management, and this knowledge deficit continues to clinical practice later when they work, which is confirmed by two studies. One study reported that the level of nurses' knowledge and attitude about pain management for postoperative children was only moderate (Prasanwon, Patoomwan, & Pookboonmee, 2018). Similarly, Mayae, Wiroonpanich, and Wattanasit (2020) presented that the mean score of knowledge and attitude regarding pain management of nurses in Southern Pediatric Intensive Care Units was at a moderate level. The most frequent barrier to nurses who provide pediatric postoperative pain management is the difficulties of accurate pain assessment, especially their lack of knowledge and skill in using and selecting the assessment tool,

leading to their inability to differentiate children's pain behaviors and understand how to use pain scales. Therefore, nurses use their own experience in assessing pain in children rather than using pain assessment tools. The authors also suggested that providing regular training to nurses about the selection of appropriate pain assessment instruments to the age of each patient is needed. The researchers suggested that workshop training in the areas they do not have enough knowledge, especially assessing pain and pediatric pain scale usage, should be provided to enhance nurses' knowledge and skills when they work in the real situation later.

Our previous work has shown that the modified FLACC scale and the COMFORT-B scale have good construct and convergent validities and both scales could measure pain in intubated and ventilated children in PICU (Narktim, Wiroonpanich, & Wattanasit, 2018). However, these scales have not been tested for inter-rater reliability between a standardized person and nurses. Therefore, this study was conducted with 2 following objectives.

Research objectives

1. The primary objective was to evaluate the inter-rater reliability of the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale between a standardized person and nurses in a Pediatric Intensive Care Unit (PICU) after the nurses had received education and skill training in assessing post-surgical pain of ventilated Thai children

2. The secondary objective was to evaluate the feasibility of using the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale in assessing post-surgical pain for ventilated Thai children in PICU.

Conceptual framework

The conceptual framework of this study was based on a KAP model (Schwartz, 1975). This model stated that the link between knowledge and attitude led to practice, especially in specific situations. In the current study, the knowledge of the pain scales with promoting attitude on using the pain scale was employed during education and skill training workshops in order to enhance the reliability of using the COMFORT-B (Thai) and modified FLACC (Thai) scales.

Methods

Research design: This research and developmental design were conducted in a PICU in Southern Thailand.

Samples: This study was conducted at the PICU of Hatyai Hospital in Songkhla province, Thailand. A total of 19 nurses participated. Inclusion criteria were 1) being a registered nurse working in a PICU, and 2) having experience of at least one year in caring for intubated children with post-surgical pain.

Research instruments: The instruments consisted of two parts: the instruments for data collection and intervention. The instruments for data collection comprised 1) a questionnaire about demographic characteristics of nurses, 2) COMFORT-B (Thai)

and modified FLACC (Thai) scoring records, 3) a questionnaire about the feasibility of using the COMFORT-B (Thai) and the modified FLACC (Thai) scales, and 4) an open-ended questionnaire for asking about problems and barriers using these two instruments.

The COMFORT-B scale was modified from the COMFORT scale originally developed by Ambuel et al. (1992) for measuring distress in ventilated children admitted to a PICU. The COMFORT-B scale can be used to assess pain in both non-ventilated and ventilated children in a PICU (van Dijk et al., 2002). It consists of 6 behavioral categories with descriptions of each item: alertness, calmness/agitation, respiratory response (ventilated children) or cry (non-ventilated children), physical movement, muscle tone, and facial tension. Each category is scored on a 1 - 5 scale. The theoretical total score ranges between 6 and 30.

The FLACC scale was originally developed by Merkel and colleagues to measure pain in children aged 2 months to 7 years who had undergone a variety of elective surgical procedures in a post-anesthesia care unit (PACU) after completing surgery (Merkel, Voepel-Lewis, Shayevitz, & Malviya, 1997). For measuring pain in intubated and ventilated children in a PICU, the item “cry” was modified to “cry face” (i.e. facial expression of moaning or crying), and the scale was named “The modified FLACC scale” (Johansson & Kokinsky, 2009). The modified FLACC scale can be used to assess pain in both non-ventilated and ventilated children. Each category is scored

on a 0 - 2 scale. The total score ranges between 0 and 10.

The COMFORT-B (Thai) and the modified FLACC (Thai) scales had been constructed by first translating the originals into Thai versions and then converting them back to English by a professional translator. The validity and reliability of the Thai versions have been previously published (Narktim et al., 2018). The results revealed that the COMFORT-B (Thai) and the modified FLACC (Thai) scales showed good construct validity when both scales discriminated pain before and during suctioning ($t = -30.98, p < .001$ and $t = -28.06, p < .001$, respectively). In addition, the inter-rater reliabilities of both scales were equally excellent (ICC = .98).

The questionnaires about the feasibilities of using the COMFORT-B (Thai) and the modified FLACC (Thai) scales were modified from the feasibility and clinical utility of the Critical-Care Pain Observation Tool (Gélinas, 2010). It consisted of 5 questions on a Likert scale, scored as 1 = not at all; 2 = a little; 3 = moderate; 4 = very much. Items included 1) The length of time to train how to use the scale is sufficient, 2) the items in the scale are clear and easy to understand, 3) the assessment of post-surgical pain in ventilated children with the scale is quick, 4) the scale is easy to use to assess post-surgical pain in ventilated children, and 5) the scale is acceptable to be used for the communication about pain among health personnel. The content validities of the instruments for data collection were

confirmed by 2 experts who work in the area of pain.

The instruments for intervention were two sets of videos that had been developed in a previous study (Narktim, et al., 2018). Each set of videos consisted of 10 videos, each taking about 2 minutes to watch.

Ethics: The study was approved by the Research Ethics Committee of Hatyai Hospital (REC-HY) (Protocol number 24/2561, dated 11/4/2561), and informed consent was obtained from all nurses who participated in this study.

Data collections: The study comprised two sessions. The first session was a workshop training providing pediatric pain knowledge, the COMFORT-B (Thai) and the modified FLACC (Thai) pain scales, and how to assess pain. The training of these two scales included a discussion of each category and item, clarification of specific behavior definitions, scoring, and showing a short video clip, followed by practicing in assessing children's pain from videos (with permission from the developer, Monique van Dijk) as similar as the real situation with post-surgical pain of intubated and ventilated children in a PICU with feedbacks until participated nurses felt comfortable and confident to use these scales.

The second session involved the testing of inter-rater reliability between the researcher as a standardized person and nurses using a set of videos. The researcher, who assumed to be the role of a standardized person, had been trained and tested for inter-rater reliability of the FLACC scale with Sandra Merkel and the

COMFORT-B scale with Monique van Dijk, yielding values of 1.00 and .95, respectively. In this study, there were two sets of videos that had been developed in a previous study (Narktim, et al., 2018). A group of nurses (3-5 nurses) was assigned to observe and independently evaluate the child's behaviors from the first set that had 10 videos using the first scale and then another scale. The order of the scales was randomized. After observing and evaluating 10 videos using both scoring scales, the agreement of each nurse with the standardized person was scored. If the proportion of percentage agreement was less than .80 (80%), the nurse of interest was required to undergo a second round of observation and evaluation again and finished with the agreement assessment to find the new proportion. As before, if a nurse scored <0.8 agreement on this round, she was required to undergo subsequent rounds, alternating between two sets of videos until she obtained an agreement proportion of at least 0.8 (80%).

After all nurses had achieved an agreement score of at least .80, the evaluation from the final round of each nurse was used to estimate the Cohen-Kappa Coefficient and 95% confidence interval for inter-rater reliability (Sheskin, 2011).

Data analysis: The personal data and the feasibility score were analyzed using descriptive statistics. The inter-rater agreement between the researcher and each nurse was analyzed using the percentage agreement formula. To calculate the percentage of agreement, the

total number of times the abstractors agreed on the same item was divided by the total number of data items. Inter-rater reliability (IRR) of the total scores was calculated by a software package using the Cohen Kappa coefficient. IRR was a measurement of the internal consistency reliability of a research tool. Using IRR required 2 or more raters to evaluate the reliability of a research tool. The Kappa value was interpreted using Altman's Benchmark scale, which was scored as < .20 = poor; .21 to .40 = fair; .41 to .60 = moderate; .61 to .80 = good; .81 to 1.00 = very good (Gwet, 2014).

Results

Nurse participants

All nurses were female with a mean age of 36.71 (SD = 8.10) years. Most of them were Buddhist (79%) and 21% were Muslim. The majority (17) had a bachelor's degree and 2 nurses had a master's degree in nursing. Years of working in the PICU varied from 1 to 19 years ($M = 11.33$, $SD = 5.73$). Thirteen nurses had received pain education 1-2 times from training sessions and six nurses had never received pain education. All of them had experience in using the modified FLACC (Thai) scale, while no one had experience in using the COMFORT-B (Thai) scale.

Inter-rater reliability of the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale

Overall, all nurses could reach the acceptable proportion agreement by the

fourth round of the evaluation using the COMFORT-B (Thai) scale. Meanwhile, it required three rounds at most when the

nurses utilized the modified FLACC (Thai) scale (Table 1).

Table 1 Numbers and percents of nurses who achieved acceptable proportion of agreement when implementing the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale ($n = 19$), separated by rounds

Scales	1st round n /number remaining (%)	2nd round n /number remaining (%)	3rd round n /number remaining (%)	4th round n /number remaining (%)
COMFORT-B (Thai)	0 (19) -	11 (19) (58%)	6 (8) (75%)	2 (2) (100%)
Modified FLACC (Thai)	4 (19) (21%)	13(15) (87%)	2 (2) (100%)	-

For the COMFORT-B (Thai) scale, the Kappa scores showed that the inter-rater reliability value for the total score was .86, which can be categorized as very good, or indicated that it is acceptable. For each item, their Kappa scores ranged from .88 to .95, which can be categorized as very good as shown in Table 2.

For the modified FLACC (Thai) scale, the Kappa scores showed that the inter-rater reliability value for the total score was .89, which can be categorized as very good or indicated that it is acceptable. For each item, the Kappa scores ranged from .90 to .98, which can be categorized as very good as shown in Table 3.

Table 2 Inter-rater reliability between a standardized person and all nurses of the COMFORT-B (Thai) scale

COMFORT-B (Thai) Category	K_{κ} [95% confidence interval]
Alertness	.95 [.90, .99]
Calmness/Agitation	.88 [.83, .93]
Respiratory response (ventilated children)	.91 [.87, .95]
Physical movement	.92 [.87, .97]
Muscle tone	.93 [.88, .98]
Facial tension	.91 [.86, .96]
Total score of the COMFORT-B (Thai) scale	.86 [.81, .91]

K_{κ} = Kappa statistic

Table 3 Inter-rater reliability between a standardized person and all nurses of the modified FLACC (Thai) scale

Modified FLACC (Thai) Category	K _κ [95% confidence interval]
Face	.90 [.94, 1.00]
Legs	.93 [.88, .97]
Activity	.91 [.76, .97]
Cry (ventilated children)	.98 [.96, 1.00]
Consolability	.97 [.94, 1.00]
Total score of the modified FLACC scale	.89 [.84, .94]

K_κ=Kappa statistic

The feasibility of using the COMFORT-B (Thai) scale and the FLACC (Thai) scale

The percentage of nurses' agreement on the feasibility of using the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale is shown in Tables 4 and 5.

For the feasibility of using the COMFORT-B (Thai) scale, 94-100% of nurses scored 4 (very much) or 3 (moderate) in 3 items. These were "The length of time used to train

the COMFORT-B (Thai) scale is sufficient", "The items are clear and easy to understand", and "The scale is acceptable to be used for the communication about pain among health personnel". Less than 80% of nurses scored 4 (very much) or 3 (moderate) in the items of "It is quick to assess post-surgical pain in ventilated children", and "The scale is easy to use to assess post-surgical pain in ventilated children." (Table 4)

Table 4 Nurses' agreements on the questionnaire about the feasibility of using the COMFORT-B (Thai) scale (*n* = 19)

Variables	Frequency (<i>n</i>)				% of nurses' agreements, which were 3 and 4
	4 (Very much)	3 (Moderate)	2 (A little)	1 (Not at all)	
1. The length of time to train in the use of the COMFORT-B (Thai) scale is sufficient.	9	10	-	-	100
2. The Items in the COMFORT-B (Thai) scale are clear and easy to understand.	4	14	1	-	94.7
3. Assessment of post-surgical pain in ventilated children with the COMFORT-B (Thai) is quick.	-	3	6	-	68.4
4. The COMFORT-B (Thai) scale is easy to use to assess post-surgical pain in ventilated children.	3	11	5	-	73.7

Variables	Frequency (n)				% of nurses' agreements, which were 3 and 4
	4 (Very much)	3 (Moderate)	2 (A little)	1 (Not at all)	
5. The scale is acceptable to be used for the communication about pain among health personnel.	5	13	1	-	94.7

For the feasibility of using the modified FLACC (Thai) scale, all nurses rated 4 (very much) or 3 (moderate) in all items, demonstrating good feasibility of the modified FLACC (Thai) scale. It included "The length of time to train the use of the scale is sufficient," "The scale has the items that are clear and

easy to understand," "The scale is quick to use to assess post-surgical pain in ventilated children," "The scale is easy to use to assess post-surgical pain in ventilated children", and "The scale is acceptable to be used for the communication about pain among health personnel." (Table 5).

Table 5 Nurses' agreement on the questionnaire about the feasibility of using the modified FLACC (Thai) scale (n = 19)

Variables	Frequency (n)				% of nurses' Agreement 3 and 4
	4 (Very much)	3 (moderate)	2 (A little)	1 (Not at all)	
1. The length of time to train in the use of the modified FLACC (Thai) scale is sufficient.	10	9	-	-	100
2. The Items in the modified FLACC (Thai) scale are clear and easy to understand.	18	1	-	-	100
3. Assessment of post-surgical pain in ventilated children with the modified FLACC (Thai) scale is quick.	16	3	-	-	100
4. The modified FLACC (Thai) scale is easy to use to assess post-surgical pain in ventilated children.	16	3	-	-	100
5. The scale is acceptable to be used for the communication about pain among health personnel.	14	5	-	-	100

Discussion

The result showed that the overall inter-rater reliability of the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale between the standardized person

and all nurses was very good as shown by the Kappa scores equaled to .86 and .89, respectively. For each item, Kappa scores ranged from .88 to .95 for the COMFORT-B (Thai) scale and .91 to .98 for the modified

FLACC (Thai) scale, which can be categorized as very good. This indicates that all nurses in the study could rate pain close to the standardized person, which in turn suggested that teaching and workshop training on how to use these scales provided adequate knowledge and skills for all nurses. Interestingly, for the COMFORT-B (Thai) scale, no nurses could pass the acceptable percent agreement ($\geq 80\%$) from the first set of 10 videos, but finally, all of them could pass by the fourth round. This finding is not surprising because all nurses in this study reported that they had never used the COMFORT-B (Thai) scale. Comparing with the modified FLACC (Thai) scale, all nurses surprisingly reported that they had experience in using the modified FLACC (Thai) scale, but only 4 nurses could pass from the first set of 10 videos. They passed by the third set of 10 videos. This meant that nurses in a PICU understand pain behaviors in each item of the FLACC (Thai) scale differently although they reported that they had experience in using the FLACC (Thai) scale. This issue points to the fact that some nurses might over or underestimate children's pain in their previous work, which means that some children might not get enough or too much pain medication. Thus, learning only pain theory by reading or training is not enough for nurses to build their competency in using a pain scale. Nurses need to be trained not only in terms of knowledge provision but also by in-service training on how to use the scale with the real patient and real

situation (Hossain, Wiroonpanich, & Orapiriyakul, 2010; Michaels et al., 2007).

Consistent with prior reports, this study demonstrated that providing pediatric pain knowledge and skills in assessing post-surgical pain in intubated and ventilated children from videos, as well as in the real clinical situation until the nurses feel comfortable and confident to use the scale with feedback, can increase their effectiveness of pain assessment (Lellan, 2004; Morrison et al., 2006). More importantly, this study suggests that it would be beneficial to pediatric pain assessment if nurses could do inter-rater reliability with the standardized person until they pass the acceptable percent agreement ($\geq 80\%$) to make sure that they have skills in using the pain scales before the real usage in the clinical practice. In addition, pediatric pain knowledge, particularly on-ward training skills using the pain scales, should be added to the nursing curriculum and allocated more than 2-3 hours.

Nurses agreed that the modified FLACC (Thai) scale had items that were clear and easy to understand, quick and easy to use to assess post-surgical pain in ventilated children, and acceptable to be used for the communication about pain among health personnel. Nurses stated that the items in the modified FLACC (Thai) scale had 5 categories and each category has only 3 items, which was easy to remember and use in clinical practice. In contrast, less than 70% of nurses agreed with the feasibility of the COMFORT-B (Thai) in terms of the items in the scale being quick and

easy to assess, which was confirmed by the report of nurses from focus group discussion. Nurses pointed out that the COMFORT-B (Thai) scale had 6 categories and each category had 5 items, which were hard to remember and took time to assess and score. Specifically, the physical movement category was really hard to assess and took time. They had to carefully observe the movement of the body and extremities and count the frequency of movement at the same time. Thus, nurses made more mistakes in the physical movement category than in other categories. However, nurses reported that the COMFORT-B (Thai) scale was a very good scale to measure pain in intubated children in a PICU because each category in this scale had 5 items showing different pain behaviors. In addition, the developers provided operational definitions for each indicator, and scores were clearly described. Thus, it was easy for nurses to understand. Nurses claimed that if they often used this scale in everyday practice until they became familiar, they may like to use this scale because the items indicated pain in children.

Limitation: The testing of inter-rater reliability between the standardized person and nurses with a set of 10 videos took time. Each video took about 2 minutes to watch. Thus, this might affect nurses' concentration. Secondly, nurses might not be able to listen to children's voices clearly from videos. This study suggests that listening to children's voices should be done in real situations.

Recommendation: The inter-rater reliability of the COMFORT-B (Thai) scale and the modified FLACC (Thai) scale between the standardized person and all nurses is a very good indicator showing that all nurses in this study could rate pain similarly to the standardized person. This study suggests that feasibility is an essential characteristic to consider when it comes to selecting a pain assessment tool. Furthermore, if a director of a hospital wants to improve pain assessment in intubated children, they should give priority to setting up workshop training incorporating inter-rater reliability for nurses.

Acknowledgments: The authors thank all PICU nurses for their diligent participation until the completion of this research. Special thanks are expressed to Drs. Monique van Dijk and Sandra Merkel, who trained the researcher to use the COMFORT-B and the FLACC scales. This research was funded by the Faculty of Nursing, Prince of Songkla University.

References

- Ambuel, B., Hamlett, K. W., Marx, C. M., & Blumer, J. L. (1992). Assessing distress in pediatric intensive care environment: The COMFORT scale. *Journal of Pediatric Psychology, 17*(1), 95-109.
- Bai, J., Hsu, L., Tang, Y., & Dijk, M. V. (2012). Validation of the COMFORT behavior scale and the FLACC scale for pain assessment in Chinese children after cardiac surgery. *Pain Management Nursing, 13*(1), 18-26. <https://doi.org/10.1016/j.pmn.2010.07.002>
- Blakely, W. R., & Page, G., G. (2001). Pathophysiology of pain in critically ill patients. *Critical Care Nursing Clinics of North America, 13*(2), 167-179. [https://doi.org/10.1016/s0899-5885\(18\)30047-9](https://doi.org/10.1016/s0899-5885(18)30047-9)
- Gélinas, C. (2010). Nurses' evaluations of the feasibility and the clinical utility of the critical-care pain observation tool. *Pain Management Nursing, 11*(2), 115-125.
- Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definition guide to measuring the extent of agreement among raters (4th ed.)*. Retrieved from http://www.agreestat.com/book4/9780970806284_chap6.pdf

- Hamill-Ruth, R. J., & Marohn, M. L. (1999). Evaluation of pain in the critically ill patient. *Critical Care Clinics*, 15(1), 35-54. [https://doi.org/10.1016/s0749-0704\(05\)70038-5](https://doi.org/10.1016/s0749-0704(05)70038-5)
- Hossain, S., Wiroonpanich, W., & Orapiriyakul, R. (2010). *Nurses' knowledge and attitude, and pain management practice of post-operative children in Bangladesh* (Mater of nursing science program, Pediatric Nursing, International Program). Faculty of Nursing, Prince of Songkla University, Songkhla, Thailand.
- Hurley-Wallace, A., Wood, C., Franck, L. S., Howard, R. F., & Liossi, C. (2019). Pediatric pain education for health care professionals. *PAIN Reports*, 4(1), e701. <https://doi.org/10.1097/PR9.0000000000000701>
- Ismail, A. (2016). The challenges of providing effective pain management for children in the pediatric intensive care unit. *Pain Management Nursing*, 17(6), 372-383. <https://doi.org/10.1016/j.pmn.2016.08.005>
- Johansson, M., & Kokinsky, E. (2009). The COMFORT behavioral scale and the modified FLACC scale in pediatric intensive care. *Nursing in Critical Care*, 14(3), 122-130. <https://doi.org/10.1111/j.1478-5153.2009.00323.x>
- Kwekkeboom, K. L., & Herr, K. (2001). Assessment of pain in the critically ill. *Critical Care Nursing Clinics of North America*, 13(2), 181-194. [https://doi.org/10.1016/S0899-5885\(18\)30048-0](https://doi.org/10.1016/S0899-5885(18)30048-0)
- Lellan, K. M. (2004). Postoperative pain: Strategy for improving patient experiences. *Journal of Advanced Nursing*, 46(2), 179-185. <http://doi.org/10.1111/j.1365-2648.2003.02977.x>
- Linton, S. L., & Shaw, W. S. (2011). Impact of psychological factors in the experience of pain. *Physical Therapy*, 91(5), 700-711. <https://doi.org/10.2522/ptj.20100330>
- Manworren, R. C. B., & Stinson, J. (2016). Pediatric pain measurement, assessment, and evaluation. *Seminars in Pediatric Neurology*, 23(3), 189-200. <http://dx.doi.org/10.1016/j.spen.2016.10.001>
- Mayae, H., Wiroonpanich, W., & Wattanasit, P. (2020). Relationship between knowledge and attitude and pain management of nurses in Southern Pediatric Intensive Care Unit. *The Southern College Network Journal of Nursing and Public Health*, 7(3), 47-57.
- Merkel, S. I., Voepel-Lewis, T., Shayevitz, J. R., & Malviya, S. (1997). The FLACC: A behavioral scale for scoring postoperative pain in young children. *Pediatric Nursing*, 23(3), 293-297.
- Michaels, T. K., Hubbart, E., Carroll, S. A., & Hudson-Barr, D. (2007). Evaluating an educational approach to improve pain assessment in hospitalized patients. *Journal of Nursing Care Quality*, 22(3), 260-265. <http://dx.doi.org/10.1097/01.NCQ.0000277784.14310.66>
- Morrison, R. S., Meier, D. E., Fischberg, D., Moore, C., Degenholtz, H., Litke, A., Siu, A. L. (2006). Improving the management of pain in hospitalized adults. *Archives of Internal Medicine*, 166(9), 1033-1039. <http://dx.doi.org/10.1001/archinte.166.9.1033>
- Narktim, K., Wiroonpanich, W., & Wattanasit, P. (2018). Comparison of the properties of pain assessment tools between the COMFORT-B scale and the modified FLACC scale in ventilated sedated children during suctioning. *Songklanagarind Journal of Nursing*, 38(1), 90-102.
- Oakes, L. L. (2011). *Compact clinical guide to infant and children pain management*. New York, NY: Springer.
- Prasanwon, K., Patoomwan, A., & Pookboonmee, R. (2018). Nurses' knowledge, attitude, and nursing practices for pain management for postoperative children. *Ramathibodi Nursing Journal*, 24(1), 37-50.
- Puntillo, K. A., Miaskowski, C., Kehrl, S., Gleeson, S., & Nye, P. (1997). Relationship between behavioral and psychological indicators of pain, critical care patients' self-reports of pain, and opioid administration. *Critical Care Medicine*, 25(7), 1159-1166. <http://dx.doi.org/10.1097/00003246-199707000-00017>
- Ramelet, A-S., Abu-Saad, H. H., Rees, N., & McDonald, S. (2004). The challenges of pain measurements in critically ill young children: A comprehensive review. *Australian Critical Care*, 17(1), 33-45.
- Schwartz, N. E. (1975). Nutritional knowledge, attitude and practice of high school graduated. *Journal of the American Dietetic Association*, 66(1), 28-31.
- Sheskin, D. J. (2011). *Handbook of parametric and nonparametric statistic procedures* (5th ed.). New York, NY, Washington: Taylor and Francis Group.
- van Dijk, M. V., Bouwmeester, N. J., Duivenvoorden, H. J., Koot, H. M., Tibboel, D., Passchier, J., & de Boer, J. B. (2002). Efficacy of continuous versus intermittent morphine administration after major surgery in 0-3-year-old infants: A double-blind randomized controlled trial. *Pain*, 98, 305-313.
- van Dijk, M. V., de Boer, J. B., Koot, H. M., Tibboel, D., Passchier, J., & Duivenvoorden, H. J. (2000). The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. *Pain*, 84, 367-377.
- Voepel-Lewis, T., Merkel, S., Tait, A. R., Trzcinka, A., & Malviya, S. (2002). 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*, 95(5), 1224-1229.
- Watt-Watson, J. H. (1997). *Relationship between nurses' empathic responses and pain management in acute care*. Retrieved from <http://www.collectionscanada.ca/obj/s4/f2/dsk2/ftp02/NQ28078.pdf>