CHAPTER 4

Results and Discussion

This chapter presents results of the study in three phases including construction of initial scale, psychometric testing, and clinical utility evaluation of the clinical pain scale for preterm neonates in the NICU.

4.1 Results

4.1.1 Phase I Construction of initial scale.

The pain indicators and scoring of the clinical pain scale for preterm neonates in the NICU was developed step by step based on foundation and linkage to the pain pathway. The results of each of the four steps are described as follows:

1) Step 1 Analyzing pain concept in preterm neonates.

Concept of pain in preterm neonates was reviewed and analyzed from empirical evidences and scholarly papers. Attributes of pain in preterm neonates are as follows:

1.1) It is an individual unpleasant, uncomfortable, and distressful experiences that originates from physical sensation from both noxious and non-noxious stimuli (indirect way).

1.2) Variables of neurophysiological and neurobehavioral reactivity respond to noxious stimulation because of anatomical and functional transition development

1.3) The developmental plasticity of nociceptive pathways includes large and overlapping peripheral receptor fields, transmitted primarily through the C-fibers and the A-delta fibers, immature descending inhibition, and lacking descending inhibitory neurotransmitters until 6 to 8 weeks after birth (dopamine, norepinephrin, and serotonin). Behavioral and physiological changes including facial expressions, transition of sleep-wakes states, and an increase in heart rate manifest as non-verbal language.

1.4) The quality of expressions are different among infants of various gestational age. The latency of pain response ranges from 2 to 30 seconds depending on the gestational age of the infant which relates to the level of neuronal and glial proliferation. Infants below 32 weeks' postmenstrual age have a significantly longer latency to change in facial expression than older infants.

Antecedents can be identified as a personal and environmental factors. Personal factors affecting pain reactivity of pain in preterm neonates include maturation, infant states, length of NICU stay, number of previous pain exposures, and respiratory support. Environmental factors are device used for noxious stimulation and implementation of cluster care in the NICU. Consequences of pain in preterm neonates compose of subsequent events and their health outcomes. Preterm neonates may or may not be provided pain relief. They cannot relief pain by him/herself, but by health care providers or parents. The infant's pain would remain the same, decrease, or increase. Altered pain and sensation processing such as primary and secondary hyperalgesia, allodynia, and windup phenomena due to early and repetitive painful stimuli. Lower body weight and head circumference, and increase intracranial pressure were found.

Within defining empirical referents, pain indicators of preterm neonates were identified according to pain pathways. The ascending pain fiber connecting with the reticular activating system and the periaqueductal gray area trigger protective autonomic responses and facial responses. Pain stimulation of the autonomic nervous system produces increased sympathetic tone and decreased parasympathetic tone. Therefore, empirical referents include facial expressions, sleep-wake state change, heart rate change and factors affecting pain reactivity. Description of pain reactivity indicators are described in three dimensions as follows:

Behavioral indicators

1) Facial expression

Four facial expressions which are the most sensitive indicator of pain in preterm neonates include brow bulge, eye squeeze, nasolabial furrow, and vertical mouth stretch. Brow bulge is bulging, creasing and vertical furrows above and between brows occurring as a result of the lowering and drawing together of the eyebrows. Eye squeeze is identified by the squeezing or bulging of the eyelids. Bulging of the fatty pads about the infant's eyes is pronounced. Nasolabial furrow is primarily manifested by the pulling upwards and furrow deepening of the naso-labial furrow (a line or wrinkle that begins adjacent to the nostril wings and runs down and outward beyond the lip corners). Vertical stretch mouth is characterized by tautness at the lip corners (vertical) coupled with a pronounced downward pull of the jaw.

2) Sleep-wake state

Sleep-wake state can be simply classified into three stages: quiet sleep, active sleep, and waking stages. In quiet sleep stage, the infant's eyes are closed, and the body is relatively motionless, with an occasional startle. In active sleep stage, the infant's eyes are closed, and they show rapid eye movement sleep or REMs (the eyes sometimes open during intense REMs); there are smiles or grimaces; the extremities and limbs show sporadic twitching; and there are occasional large movements of the arms and legs with moderate positional displacement within the incubator. In waking stage, the infant's eyes may be closed during very active periods with obvious fussing or crying, accompanied by large movements or flailing of the arms and legs, and the eyes could be open during periods of quiet wakefulness.

Physiological indicator: heart rate change

Heart rate change is the different number between two values of heart rate. An increased heart rate is that the later heart rate is faster than the previous heart rate.

Factors affecting pain reactivity

Factors affecting pain reactivity of preterm neonates included gestational age at birth, respiratory support, number of previous pain exposures, and length of NICU stay. In terms of gestational age at birth, preterm neonates with less than 32 weeks of gestation, showed dampened behavioral response to painful stimuli (Gibbins, Stevens, Beyene et al., 2008) and their sleep-wake states were increasingly distinct after 32 weeks (Foreman et al., 2008). In terms of the length of NICU stay and receiving respiratory support such as mechanical ventilation were factors related to frequency of painful procedures influencing pain reactivity of preterm neonates (Cruz, Fernandes, & Oliveira, 2015; Williams, Khattak, Garza, & Lasky, 2009). Preterm neonates receiving nasal oxygen, continuous positive airway pressure and ventilation support were exposed to a high number of painful procedures (Cruz et al., 2015). In addition, one study stated that preterm neonates who received more than 20 procedures were associated with dampened facial activity measured by NFCS (Grunau et al., 2001).

Information from the concept analysis and literature review of the pain pathway clearly guides boundary of pain in preterm neonates and shows that it is multi-dimension phenomenon. Pain in preterm neonates is defined as an acute unpleasant sensory and emotional experience associated with actual or potential tissue damage caused by medical or nursing procedures that invade the preterm neonate's body integrity, causing skin injury or mucosal injury. Transmitting this signal involves with nociceptors in the periphery and conducts signals to the dorsal root ganglion in the spinal cord through C-fiber. The releasing inflammatory soup stimulates signal processing to the brain, then perception of signal. To measure procedural pain in preterm neonates has to detect from behavioral, physiological, and factors affecting pain reactivity. Information from the first step also led to the draft list of pain indicators and its description which included increased four facial expressions, sleep-wake states, heart rate, and factors affecting pain reactivity such as gestational age at birth, type of respiratory support, number of previous pain exposures, and length of NICU stay were constructed for the structured observation checklist of pain indicators for the next step.

2) Step 2 Generating a list of pain indicators by clinical observations

All fifteen painful procedures which were observed at the bedside in NICUs were blood sample collections including twelve heel sticks by the same phlebotomists and three venipuncture procedures by the same nurse. The duration of the puncture phase for 15 occasions ranged from 34 seconds to ten minutes and 12 seconds, with a mean value of two minutes and 51 seconds (a mean value of one minute and 38 seconds for heel stick and six minutes 54 seconds for venipuncture). Information and the frequencies of occurrence of each pain indicator during the baseline, puncture, and recovery phase were described as follows:

2.1) Behavioral response: facial expression and sleep-wake state

Regarding the 15 occasions of painful procedures, brow bulge, eye squeeze, and nasolabial furrow were found in the baseline phase of a few occasions, whereas mouth stretch or crying was not found. During the puncture phase, the varied expressions from slight, intense, to robust grimace were observed. Most preterm neonates immediately expressed at least one of brow bulge (11 occasions), eye squeeze (ten occasions), nasolabial furrow (eight occasions), and vertical mouth stretch (five occasions) (see Figure 4.1). In the recovery phase, most preterm neonates had relaxed facial expression and all of brow bulge, eye squeeze, and nasolabial furrow remained only in one occasion.

Sleep-wake state was composed of quiet sleep, active sleep, and waking. A wide range of sleep-wake state occurred ranging from quiet sleep to waking in all three phases. During the puncture phase, most infants showed a behavioral state of arousal or transition between behavioral states with being more awake relative to the baseline. The arousal responses to painful stimuli differed from tactile stimuli during the warming or antiseptic period. In addition, such behaviors were not seen in the baseline and recovery phases.

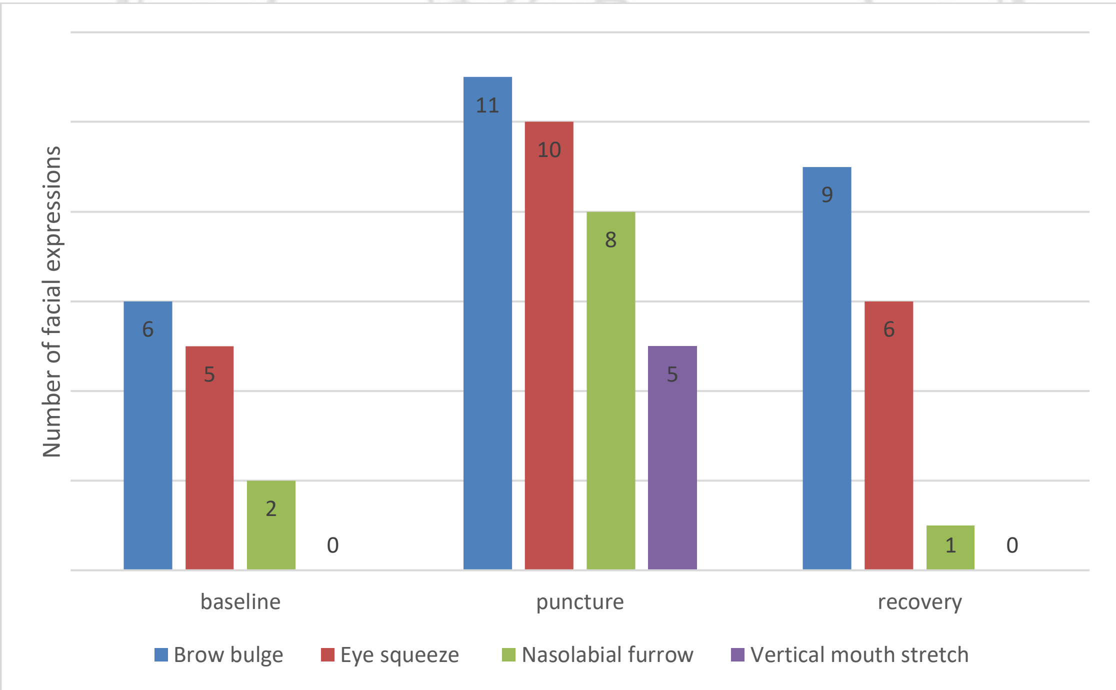


Figure 4.1 Four facial expressions in each phase of 15 painful procedure occasions

2.2) Physiological response: heart rate change

According to 15 occasions of painful procedures, the mean heart rates in the baseline, puncture, and recovery phase were 156.17, 164.77, and 158.37 beats/minute, respectively (see Table 4.1). During the baseline phase, the heart rate changed within a narrow range, 1.27 to 5.82 with the median of 2.60. The heart rate change > 5 beats/minute was found in only one occasion. Comparing with the baseline data, an increased heart rate and heart rate fluctuation were noted > 30 seconds after a needle puncture. Of 15 occasions, an increased heart rate was found on 13 occasions. The mean difference between the heart rate at the baseline and puncture phases was 7.99 beats/minute (range, -11.77 to 26.04). Comparing with the puncture data, the heart rate in the recovery phase of all infants decreased but not all infants returned to baseline values on each occasion. The mean difference between the heart rate in the puncture and recovery phases was 7.27 beats/minute (range, 0.50 to 20.07). As expected, the mean difference between the heart rates in the baseline and recovery phase was only -0.73 beats/minute (range, -11.00 to 12.27 beats/minute). Most of different values between the heart rate in the puncture phase compared with the baseline and recovery were positive values which means that the heart rates in the puncture phase were mostly higher than those in the baseline and recovery phases. Regarding notes from two observers, most preterm neonates had a rapidly increase in heart rate during needle insertion, especially 30 to 60 seconds after having needle insertion, and remained high (may or may not be tachycardia) along the time of squeezing, then slowly decreased mostly by themselves. However, the heart rate did not return to the baseline value. It can be concluded that heart rate in most occasions increased after 30 seconds of skin breaking. Since in the baseline phase, the heart rate variability no more than 5 beats per minute was mostly observed and consistent with the previous study (McIntosh et al., 1994), an increased heart rate more than five beats/minute could be used as an indicator of pain reactivity.

2.3) Factors affecting pain reactivity

Results from the personal data profile of eight preterm neonates and fifteen bedside observations revealed that the high frequency of blood sample collections was found during the first and the second week after birth (mean = 52.33 procedures, range 7 to 81). Most of the collections were for determining blood gas with some for checking glucose levels. Determining blood gas was required more frequently in preterm neonates with ETT than those with NPCPAP or cannula. Interestingly, the longer NICU stay and the more previous pain exposures, the less behavioral responses to noxious stimulation. The pain reactivity of the repeated procedure in one infant (case number six) with 28 weeks' gestational age at birth was different when compared with herself (see Table 3.2). A follow-up four observations of her pain reactivity found that behavioral responses to noxious stimulation gradually decreased from the first to the last occasions. Four factors including gestational age at birth, length of NICU stay, respiratory support, and number of previous pain exposures since birth might influence the pain reactivity of a preterm neonate. Therefore, at the end of this step there were seven indicators including facial expression, sleep-wake state, heart rate change, and four factors affecting pain reactivity. This list of seven indicators was used for the next step (see Figure 4.2).

Table 4.1

Mean of heart rate in each phase and the difference between two phases (n = 15)

No.	Mean of heart rate (beats/minute)			Difference between two phases		
	Baseline	Puncture	Recovery	Puncture-Baseline	Puncture-Recovery	Baseline-Recovery
1	170.83	172.73	169.23	1.9	3.50	1.60
2	162.63	178.80	165.03	16.17	13.77	-2.40
3	165.53	174.00	166.33	8.47	7.67	-0.80
4	168.43	169.08	165.40	0.65	3.68	3.03
5	182.09	175.00	172.29	-7.09	2.71	9.80
6	147. 05	169.50	149.43	22.45	20.07	-2.38
7	158.97	159.95	156.87	0.98	3.08	2.10
8	152.77	141.00	140.50	-11.77	0.50	12.27
9	147.57	158.58	157.10	11.01	1.48	-9.53
10	155.63	181.67	166.63	26.04	15.04	-11.00
11	158.43	175.00	165.83	16.57	9.17	-7.40
12	160.03	173.81	163.97	13.78	9.84	-3.94
13	161.40	169.22	161.77	7.82	7.45	-0.37
14	158.44	168.78	160.60	10.34	8.18	-2.16
15	152.33	154.93	152.07	2.6	2.86	0.26
Mean	156.17	164.77	158.37	7.99	7.27	-0.73

Checklist of Seven Pain Indicators

1. Facial expression
    - ❐ Brow bulge (no change, slightly, robust)
    - ❐ Eye squeeze (no change, slightly, robust)
    - ❐ Nasolabial furrow (no change, slightly, robust)
    - ❐ Vertical mouth stretch (no change, slightly, robust)
2. Sleep-wake state
    - ❐ Maintains the same state
    - ❐ Transit from one state to another
    - ❐ (Silent) Cry or tense body
3. Heart rate (HR) change
    - ❐ Stable and normal HR
    - ❐ Increased HR 5 beats/minute from baseline
    - ❐ Tachycardia with oxygen saturation drop
    - ❐ Fluctuating or increased HR > 10 beats/minute from baseline
4. Factor affecting pain reactivity: Gestational age
    - ❐ < 32 week
    - ❐ ≥ 32 to $36^{6/7}$weeks
5. Factor affecting pain reactivity: Length of NICU stay
    - ❐ 1 week
    - ❐ 2 weeks
    - ❐ > 2 weeks
6. Factor affecting pain reactivity: Respiratory support
    - ❐ Room air
    - ❐ Nasal oxygen
    - ❐ CPAP or ventilation
7. Factor affecting pain reactivity: Previous pain exposures
    - ❐ 0-20 times
    - ❐ > 20 times

Figure 4.2 Checklist of seven pain indicators derived from step 1 and 2

3) Step 3 Determining the format for measurement by clinical experts' interview

Data pertaining meaning and characteristics of pain, pain indicators, scoring, and use of pain scale in the clinical setting from five clinical experts' interviews were analyzed and summarized as follows:

3.1) Meaning and characteristics of pain in preterm neonates

The clinical experts explained pain in preterm neonates as a multi-dimensional concept. They stated that all kinds of environmental stimuli such as invasive procedures and sensory-overstimulation can cause pain and stress leading to adverse consequences. Distinguish between pain and stress was very difficult but it was possible because any skin breaking procedures in preterm neonates causes an unpleasant sensory and emotional experience or pain. They expressed that it is challenging for nurses to assess pain in such vulnerable infants, especially, in the critical first few days after birth. Unfortunately, most health care providers focus only on clinical conditions such as respiratory failure and leading causes of infant mortality.

3.2) Pain indicators in preterm neonates

The clinical experts generally agreed with the seven pain indicators. The detail of each indicator is described as follows:

Focusing on the four facial expressions, all experts agreed with using facial expression as a pain indicator in preterm neonates. They explained that facial expressions can be seen in preterm neonates with > 28 weeks of gestational age and are more clearly observed in preterm neonates with > 32 weeks of gestational age. For optimal and practical use in clinic, some experts suggested that facial assessment should be divided into two parts including upper facial expression (brow bulge and eye squeeze) and lower facial expression (nasolabial furrow and mouth stretch). The clinical experts made statements as follows:

"*…facial expression showed first such as brow bulge…the more they had pain, their experienced facial expressions were more clearly observed…*"

"*Even though his/her face was so tiny and strapped by plaster with ETT, it was possible to grade the score on face activity. But it would be nice and easier to assess* *each part separately. I think using upper and lower facial expressions would be optimal for clinical observations at least for me*"

Focusing on the transition among three sleep-wake states, all experts positively agreed with using sleep-wake states as a pain indicator in preterm neonates. They explained that an infant begins more arousal during antiseptic phase and was more awake during the puncture phase. Some experts noticed that no change of arousal has occurred in a weakened infant such as ELGA infants or infants with repeated painful exposures. The clinical experts shared their perceptions as follows:

"*…within the first 48 hours of life, the tiny infants such as 24 weeks' gestational age won't cry. They were too fragile and too tiny. They lay still with no muscle tone."*

"*I agree with changing from sleep to wake state or from light sleep to cry but I'm not sure in infants with 23-25 gestational age or a preterm infant with 500-600 grams because they have no energy to cry or even open their mouths.*"

"Because of the endotracheal tube with plaster strapping, it is quite hard to observe their crying. But if you are an NICU nurse, it can be seen with careful observation. You will see their faces look like silent crying and changes from one state to another during the procedure."

Focusing on the heart rate change, the clinical experts stated that an increase in or fluctuation of heart rate and a decrease in oxygen saturation occurs during painful events. For the severely painful procedure, those changes were more clearly observed. Heart rate changes are more indication of pain response than oxygen saturation changes do. However, infants may have tachycardia or oxygen saturation drop in some conditions without painful stimulation (i.e., side effect of some medications, severe illness condition, and just having some nursing interventions). Thus, heart rate change might be the better pain indicator than tachycardia. Due to the normal heart rate reference value of preterm neonate in general, the normal range of 120-160 beats/minute was widely used in NICU. Adding such values as normal heart rate or tachycardia instead of using words might be more clearly understanding for the users. However, the heart rate change of preterm neonates during baseline and stable condition was never more than five beats/minute. Therefore, the experts suggested to classify heart rate in either heart rate value or heart rate change (see Table 4.2). The examples of clinical experts' statements are as follows:

"*The physiology of preterm neonates' change depends on the severity of pain. From my experience, when they encounter painful procedure, their heart rates shoot and oxygen saturation swing or may drop. On the contrary, if no change or there is stability in heart rate and oxygen saturation, it means they can rest and we can say that they feel no pain.*"

"*…their pain responses are increasing heart rate, work of breathing, and blood pressure…* "

"*…it happens and can be observed during the procedure. Their heart rate is rapidly increasing and suddenly drops after it has been done. In some cases, they might turn to blue or apneic if it was a prolonged or severe painful procedure…*"

Focusing on factors affecting pain reactivity, most clinical experts supported the idea that preterm neonates with different ages react to or have ability to respond differently. In terms of neuroanatomical maturity ready for response, it seemed practical and reasonable for them to use 32 weeks of gestational age at birth. However, the experts suggested that further literature review for the specific information of age should be performed. They explained that preterm infants can learn from the previous painful experience; therefore the length of NICU stay should be used rather than postnatal age. The infants show some cues that they know what's going on. For example, a preterm neonate responds to pick up his/her leg by avoidance behaviors such as movement and withdrawal his/her limbs. Two of the three clinical experts, one nurse and one neonatologist disagreed with using respiratory support as a pain indicator, even though they agreed with mode of respiratory support related to the number of laboratory and severity of illness. All clinical experts agreed with using the number of previous procedures as pain indicator. They agreed with scoring based on the reviewed literature and provided options how to collect that information in the clinical setting.

The clinical experts shared their experiences as follows:

"*Each preterm infant has a different threshold depending on age and pain exposures. The mature infants who have painful experiences have high sensitivity. They* *rapidly show pain reactivity. Therefore, preterm infant with no response can be interpreted that there is no pain or pain but en-durable.”*

“*If evidence from reviewed literature clearly shows that the number of previous pain procedures affects pain reactivity, NICU nurses should be aware of it. It must be one of the pain indicators. At least, it is raising the awareness of heath care providers during their assessment.*”

3.3) Scoring of each indicator

All clinical experts agreed to use 32 weeks’ gestational age as a cut-off point age for dividing behavioral expressions of preterm neonates. One clinical expert strongly supported this idea. She has been working in the hospital that uses the gestational age of ≥ 32 weeks as a criteria for the NICU admission. The inborn infant who was younger than 32 weeks’ gestational age was usually transferred from her unit to the NICUs at Chiang Mai University Hospital, therefore she has great opportunity to screen and to clearly picture the difference between those infants. The infant born with more gestational age (≥ 32 weeks) displayed more vigorous behaviors and pain reactivity than the infant born with less gestational age (< 32 weeks).

The clinical experts also provided feedback to raise weight of the physiological indicator scoring in preterm neonates, especially in infants less than 32 weeks’ gestational age. Therefore, the heart rate scoring was identified into four levels (0, 1, 2, and 3) in the preterm infants with < 32 weeks of gestational age group. Some experts suggested adding oxygen saturation drop in the scale. Therefore, the scoring of heart rate indicator was revised by adding change of oxygen saturation. For behavior indicators, the experts suggested that the scoring not be divided into many levels because it is hard to adjust or differentiate between each level of scores, especially when it is used with very low gestational age infants. In addition, behavior scoring should not be done immediately after tissue damage because most preterm neonates have a delayed response.

3.4) Use of pain scales in clinical setting

The experts were concerned with obstacles of clinical utilization such as the low level of NICU nurses’ awareness and understanding of pain assessment in preterm neonates. To raise awareness of assessors, all clinical experts agreed with a training session before the initial assessment before the scale was implemented in the clinical setting. As suggested by the experts, the core content of a training session for health care providers should consist of basic knowledge of the pain pathway, consequences of untreated pain in preterm neonates, fundamentals of each indicator, and instruction of assessment tool. In order to gain observational skills, especially for scoring behavioral indicators, a video of case scenario was suggested to be included in the training session.

The experts insisted that monitoring pain as a routine nursing care with a valid tool should be performed. By the way, the simplicity and feasibility of the pain scale was a very important influence for the frequency of usage. A pain score was recorded in the vital signs flowchart and nurse's note. Documenting the pain score has been an unresolved and continuous problem for nurse due to time pressure and work load. The clinical experts shared their experiences as follow:

"*Nurses' awareness and understanding should be the first priority for training session because it is very important to change their attitude about pain management behaviors.*"

"*If it was easy to use and simply to remember like Apgar score, everyone would use it more frequently.*"

"*All painful procedures had to be performed to help the patients during critical period. There is a rule for helping the infant which is to stop for soothing after three or more failure of punctures.*"

At the end of this step, all suggestions from clinical experts were used to draft the initial scale for the next step (see Table 4.2). Three major points for developing the initial clinical pain scale were separating upper and lower facial expressions as two indicators, using different scoring criteria for two age groups, and identifying scoring criteria of three factors affecting pain reactivity based on the previous studies.

4) Step 4 Having the initial clinical pain scale reviewed by content experts

The results of the first-round review of six experts were as follows:

4.1) The relevance of indicators to pain in preterm neonates

Six out of seven indicators were rated by the experts as relevant to a pain reactivity in preterm neonates with I-CVI values of .83 to 1.00 (Appendix F-1). Thus, these indicators were retained. Only one indicator, respiratory support, was rated as not having relevance as a pain indicator in preterm neonates with I-CVI level of .33. With respect to the experts' comments and an I-CVI of the respiratory support; therefore, it was eliminated from the scale. The S-CVI/Ave of the scale with six remaining indicators was .92. However, the experts were still concerned that the frequency of procedures predicted the types of respiratory support and higher numbers of skin breaking procedures since birth could be used to predict dampened facial responses to lancet. Therefore, the experts suggested that types of respiratory support be collected as a demographic data. Finally, six pain indicators of upper facial expression, lower facial expression, sleep-wake states, heart rate, length of NICU stay, and previous pain exposures were judged as relevant and included in the scale.

4.2) The relevance of indicator scoring to pain in preterm neonates

Scoring of the remaining six indicators in preterm neonates with < 32 weeks' gestation and ≥ 32 to $36^{6/7}$ weeks gestation were considered. Five indicator scoring items, except scoring of heart rate, were rated as relevant to pain in preterm neonates with I-CVI level of .83 to 1.00. The I-CVI values of heart rate change scoring in groups of preterm neonates with < 32 weeks and ≥ 32 to $36^{6/7}$ weeks' gestational age were .33 and .67, respectively (see Appendix F-2). The experts commented that heart rate change from baseline be used instead and only three levels for scoring (0, 1, and 2) were appropriate for bedside clinical observation. Use of the same criteria between the two groups of gestational ages was recommended. After revising the clinical pain scale based on the experts' recommendations, the second-round evaluation of the heart rate scoring was established and the results were presented in Table 4.3. The second-round review by six experts revealed values the I-CVI of 1.00 for heart rate scoring of both groups (see Appendix F-3). Thus, the revised clinical pain scale with S-CVI/Ave for scoring indicators of was .94 was used in the next step.

Table 4.2

The indicators and scoring format of the initial clinical pain scale

Indicator	Score	GA at birth	
		< 32 weeks	≥ 32 to $36^{6/7}$ weeks
Upper facial expression (brow bulge and eye squeeze)	0	Relaxed	Relaxed
	1	Frown slightly (eyes slightly closed and shallow furrows)	Frown robust
	2	Eyes tightly closed or no response to traumatic procedure, gaze avoidance deep furrows	Frown intense
Lower facial expression (nasolabial furrow and mouth stretch)	0	Relaxed	Relaxed
	1	Opening mouth slightly	Opening mouth robust
	2	Robust	Intense
Sleep-wake states	0	Maintains the same state	Maintains the same state
	1	Transit from one state to another with waking and no cry	Transit from one state to another with movement and tense body
	2	Waking and (silent) cry	Cry or tense body
Heart rate change	0	Stable baseline of 120-160 bpm	Stable baseline of 120-160 bpm
	1	Increase > 5 bpm from baseline	HR > 160 bpm with decrease oxygen saturation
	2	HR > 160 bpm	Increase > 10 bpm from baseline
	3	Increase >10 bpm from baseline or fluctuating but > 160 bpm	

Table 4.2 (continue)

Indicator	Score	GA at birth	
		< 32 weeks	≥ 32 to $36^{6/7}$ weeks
Respiratory support	0	No history of oxygen therapy and oxygen support not required; Room air	No history of oxygen therapy and oxygen support not required; Room air
	1	Having or had history of requiring Nasal oxygen	Having or had history of requiring Nasal oxygen
	2	Having or had history of requiring CPAP or ventilation	Having or had history of requiring CPAP or ventilation
Length of NICU stay	0	> 2 week	> 2 week
	1	2 weeks	2 weeks
	2	1 week	1 week
Previous pain exposures*	0	≥ 20 procedures	≥ 20 procedures
	1	0-19 procedures	0-19 procedures

Note. * Number of previous blood collection procedures since at birth (manually count from medical chart or bedside recording)

4.3) The evaluation of the scale format

The six content experts evaluated the appropriateness of scale format. The numbers experts rated in each item was summarized. The majority of the experts marked on the "quite appropriate" level in almost all items, except for items of baseline recordings (see Table 4.4). Four experts marked on the "most to quite appropriate" level in photograph of facial expression scoring (see Figures 4.3-4.4). The majority of expert marked on the "most to quite appropriate" level in using plot graph for recording the total pain score (see Figure 4.5). They suggested that clarification of when and how to assess heart rate and sleep-wake states during baseline phase should be declared. The scale format, especially baseline recording, was refined based on the experts' comments and suggestions.

Table 4.3

Heart rate scoring suggested from experts of the 1st and 2nd round

Heart rate score	Expert review	Gestational age at birth	
		< 32 weeks	≥ 32 to 36 6/7 weeks
0	1st round	Stable baseline of 120-160 bpm	Stable baseline of 120-160 bpm
	2nd round	Increase < 5 bpm from baseline	Increase < 5 bpm from baseline
1	1st round	Increase > 5 from baseline	HR > 160 bpm with decrease oxygen saturation
	2nd round	Increase ≥ 5-9 from baseline	Increase ≥ 5-9 from baseline
2	1st round	HR > 160 bpm	Increase > 10 from baseline or fluctuating but > 160 bpm
	2nd round	Increase ≥ 10 from baseline	Increase ≥ 10 from baseline
3	1st round	Increase > 10 from baseline HR or fluctuating but > 160 bpm	-
	2nd round	(deleted)	-

After the experts' review, the main refinements of the initial clinical pain scale (see Table 4.2) were eliminating a respiratory support indicator, revising format of two gestational age groups (put criteria of two age groups underneath the behavioral indicators), using the concise words for explaining each behavior indicator, identifying an increased heart rate scoring for both gestational age groups, and using unit of day rather than week in length of NICU stay.

The clinical pain scale with six indicators and refined format was used for psychometric testing in Phase II and clinical utility evaluation in Phase III. The photograph of facial expression scoring and graph were not used in the psychometric testing in Phase II, but video scenario of those same cases was used instead in training of clinical utility evaluation in Phase III.

Table 4.4

Experts' evaluation of the scale format (n = 6)

Items		Level of appropriateness			
		Most	Quite	Some what	Not appropriate
1.	Description of scale instruction	1	4	1	-
2.	Step of scale instruction	1	3	2	-
3.	Table pattern of the scale	1	4	1	-
4.	Baseline recording	2	1	3	-
5.	Scoring table	2	4	-	-
6.	Language of scoring in each indicator	1	4	1	-
7.	Photograph of facial expression scoring	1	3	2	-
8.	Graph pattern of total score	1	5	-	-
9.	Plot total score in graph	1	5	-	-
10.	Recording of nursing intervention with pain scoring	3	3	-	-

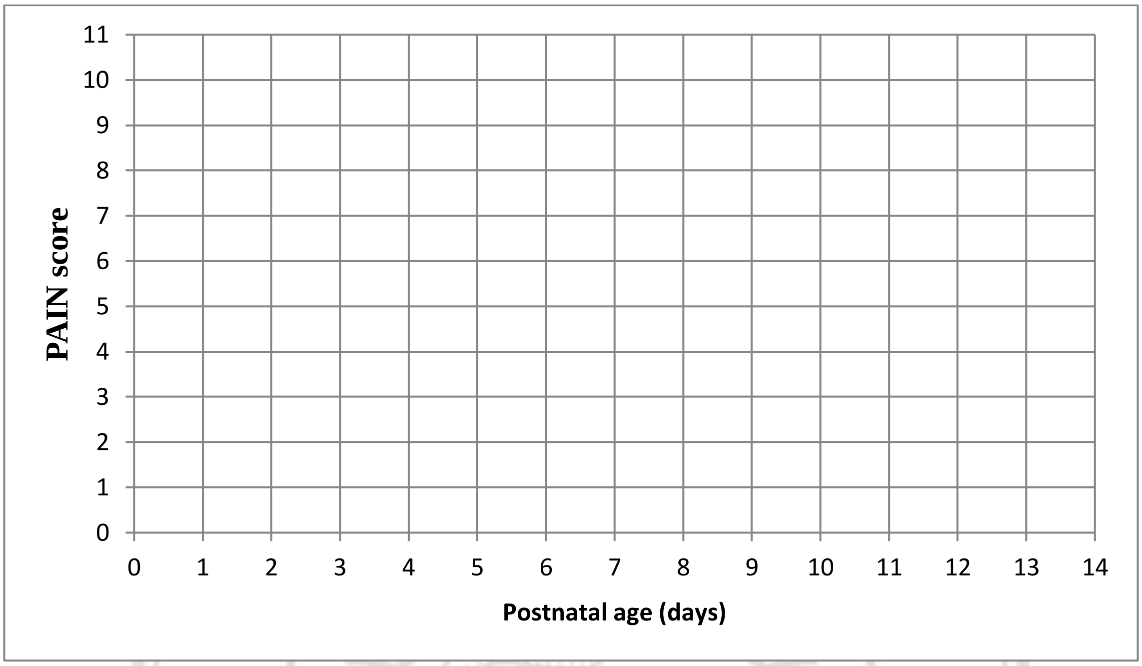
Three preterm infants (<32 weeks of gestation) with relaxed upper facial expression scored zero point and relaxed lower facial expression scored zero point

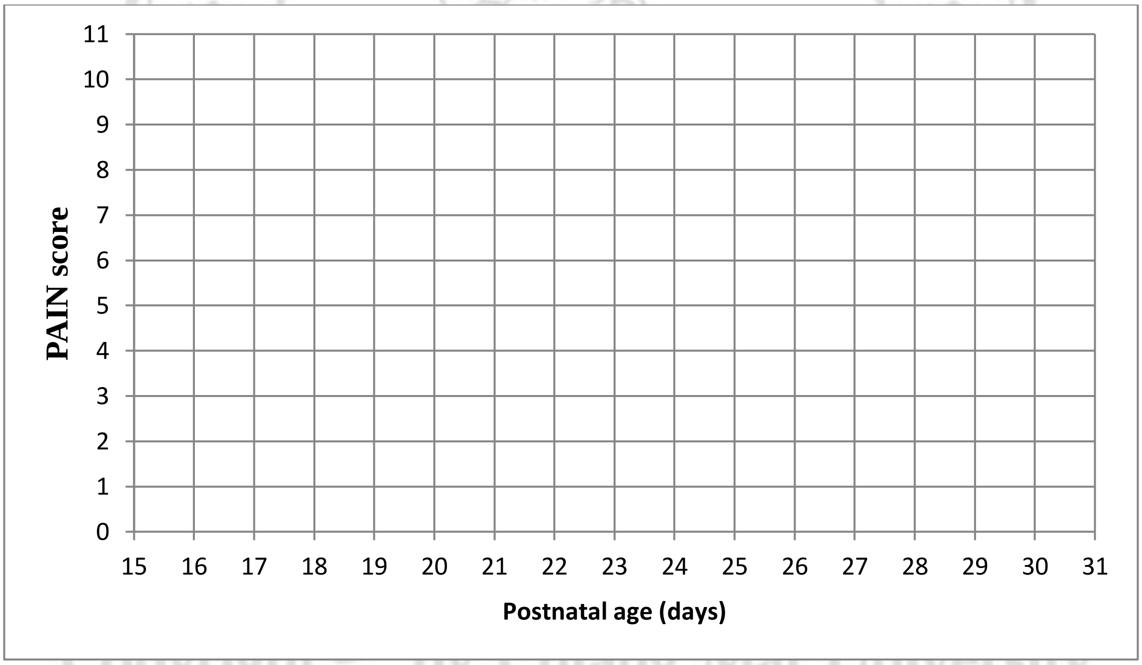
A preterm infant (< 32 weeks of gestation) with brow bulge scored one point and relaxed (left-sided), a preterm infant (< 32 weeks of gestation) with brow bulge and eye squeeze scored two points and nasolabial furrow scored one point (middle), and a preterm infant ($\geq$32 to $36^{6/7}$ weeks of gestation) with brow bulge and eye squeeze > once scored two points and nasolabial furrow and open mouth slightly scored one point (right-sided)

Figure 4.3 Example one of facial expression scoring

A preterm infant (<32 weeks of gestation) with brow bulge and eye squeeze scored two point and nasolabial furrow and open mouth widely scored two points (left-sided) and a preterm infant (<32 weeks of gestation) with brow bulge and eye squeeze scored two point and nasolabial furrow and open mouth widely scored two points (right-sided)

Figure 4.4 Example two of facial expression scoring





Instruction: Plot X on the graph and record nursing intervention with abbreviation of English alphabet as following: M = medication, S = sucrose, T= touch, B = breastfeeding or breast milk, NS = nonnutritive sucking, KC= kangaroo care, FC = facilitated tuck, SW = swaddling, and P = Parenting care

Figure 4.5 Graph for total pain score with nursing intervention with pain scoring

Table 4.5

Clinical Pain Scale for Preterm Neonates

Indicators	Findings	Points
Upper facial expressions		
GA at birth < 32 weeks	Relaxed	0
	Brow bulge	1
	Brow bulge and eye squeeze	2
GA at birth ≥ 32-$36^{6/7}$ weeks	Relaxed	0
	Brow bulge and eye squeeze once	1
	Brow bulge and eye squeeze > once	2
Lower facial expressions		
GA at birth < 32 weeks	Relaxed	0
	Nasolabial furrow	1
	Nasolabial furrow and open mouth slightly	2
GA at birth ≥ 32-$36^{6/7}$ weeks	Relaxed	0
	Nasolabial furrow and open mouth slightly	1
	Nasolabial furrow and open mouth widely	2
	Sleep-wake states#	
GA at birth < 32 weeks	No change	0
	Waking and no cry	1
	Waking and cry	2
GA at birth ≥ 32-$36^{6/7}$ weeks	No change	0
	Movement and tense body	1
	Cry	2
An increased heart rate (HR)*	< 5 beats/min from baseline	0
	≥ 5-9 beats/min from baseline	1
	≥ 10 beats/min from baseline	2
Length of NICU stay	> 14 days	0
	8-14 days	1
	0-7 days	2
Previous pain exposures	≥ 20 procedures	0
	0-19 procedures	1

Note. #State change in relation to the baseline pattern (quiet sleep, active sleep, & waking)

*An increased HR in relation to baseline HR …… beats/min and puncture HR …... beats/min (Baseline phase = 30 seconds before puncture)

4.1.2 Phase II Psychometric Testing

1) Reliability

1.1) Internal consistency

According to the inter-items correlation matrix among six indicators, there were positive and high relationships among the behavioral indicators including upper and lower facial expressions and sleep-wake state (*r* value ranged from .94 to .95) (see Table 4.6). The correlation coefficients between heart rate scores and scores of each behavioral indicator were .61 to .66 in all three phase (*n* = 159) (see Table 4.6). The positive and moderate relationship between two factors affecting pain reactivity, length of NICU stay and previous pain exposures, was found (*r* = .51), whereas those factors have low relationship with all of the rest of the indicators (range, .05 to .25). In addition, results of the items analysis indicated that these two factors had values of item-total correlation lower than .30 (see Table 4.7). Therefore, two factors affecting pain reactivity were eliminated from the clinical pain scale. The four-indicator clinical pain scale had Cronbach's alpha coefficients of .86 for during procedures (*n* = 53) and .94 for all three phases (*n* =159).

Table 4.6

Correlation matrix among six pain indicators in all three phases (n = 159)

Pain indicators	Upper facial expression	Lower facial expression	Sleep-wake states	Heart rate	Length of NICU stay
Upper facial expression	1.00				
Lower facial expression	**.94**	1.00			
Sleep-wake state	**.95**	**.94**	1.00		
Heart rate change	**.63**	**.61**	**.66**	1.00	
Length of NICU stay	.20	.25	.17	.05	1.00
Previous pain exposures	.09	.10	.06	.05	**.51**

Table 4.7

Internal consistency of the six-indicator clinical pain scale in all three phases (n = 159)

Indicators	Mean (*SD*)	Scale mean if items deleted	Corrected item-total correlation	Cronbach's Alpha if items deleted
Upper facial expression	0.78 (0.93)	4.14	.86	.74
Lower facial expression	0.72 (0.92)	4.21	.87	.74
Sleep-wake state	0.75 (0.91)	4.17	.85	.74
Heart rate change	0.91 (0.90)	4.01	.58	.81
Length of NICU Stay	1.28 (0.86)	3.64	.26	.89
Previous pain exposures	0.49 (0.50)	4.43	.20	.86

1.2) Inter-rater reliability

Mean scores (*SD*), range, and intraclass correlation coefficient (ICC) of the four-indicator clinical pain scale scores of three phases by two raters (the registered nurse and the researcher) in each of the 53 occasions were in Table 4.8. The ICC values of the four-indicator clinical pain scale ranged from .91 (for the puncture phase) to 1.00 (for the recovery phase).

2) Construct validation

The construct validity of the clinical pain scale was examined using hypothesis testing approach and convergence evidence. Kruskal-Wallis and Mann-Whitney tests were calculated for hypothesis testing. Spearman rank correlation coefficients were calculated for convergence validity.

2.1) Result of hypothesis testing

The highest mean scores of all behavioral indicators and the physiological indicator were found in the puncture phase. The mean scores of those indicators in the recovery phase were greater than those in the baseline phase (see Table 4.9). The results from 53 observed occasions of phase II in this study found that the mean heart rate during the baseline, 30 seconds, 60 seconds, the recovery phases were 158.48, 169.51, 174.45, and 165.21 beats/minute, respectively.

The mean total pain scores of baseline, puncture, and recovery phases were 0.37, 6.57, and 2.49, respectively. Kruskal-Wallis testing indicated that the median total pain scores among baseline, puncture and recovery phases were significantly different (Chi-Square 95.95, p = .000). This finding indicated that at least one pair of median scores was significantly different. The comparison of using Mann-Whitney tests indicated that median score of puncture phase was significantly different from the baseline and recovery phases (p = .000) (see Table 4.10). These results suggest that pain score in the puncture phase assessed by the clinical pain scale was significantly higher than those in the baseline and recovery phases.

2.2) Evidence of construct validity: convergence validity

Spearman rank correlation coefficients were calculated to assess the relationship between pain scores which were assessed by the newly developed clinical pain scale and PIPP-R scale. There was a positive relationship between the pain scores which were assessed by two scales in the baseline phase (r_s= .375, p = .006), puncture phase (r_s = .789, p = .000), and the recovery phase (r_s = .878, p = .000) (see Table 4.11).

Table 4.8

Inter-rater reliability of the four-indicator clinical pain scale assessed by two raters

Phase	Rater	Total Pain Score			α	ICC (95% CI)
		Mean	*SD*	Range		
Baseline	1	0.38	1.18	0-8	.99	.98 (.98 - .99)
	2	0.36	1.18	0-8		
Puncture	1	6.66	2.00	2-8	.95	.91 (.84 - .94)
	2	6.47	2.09	2-8		
Recovery	1	2.49	2.89	0-8	1.00	1.00 (. - .)
	2	2.49	2.89	0-8		

Table 4.9

Mean of each four-indicator across three phases (n = 53)

Pain indicators	Mean scores (*SD*) in each phase		
	Baseline phase	Puncture phase	Recovery phase
Upper facial expression	0.08 (0.33)	1.69 (0.57)	0.57 (0.84)
Lower facial expression	0.04 (0.27)	1.58 (0.69)	0.51 (0.87)
Sleep-wake state	0.04 (0.27)	1.66 (0.51)	0.55 (0.87)
Heart rate change	0.22 (0.52)	1.64 (0.59)	0.87 (0.90)
Total scores	0.37 (1.17)	6.57 (2.00)	2.49 (2.89)

Table 4.10

Comparison of pain scores in three phases using the clinical pain scale (n = 53)

Phase	Mean	*SD*	Mean rank	Median	*Mann-Whitney U tests* (Puncture)	*p*
Baseline	0.37	1.17	40.01	0	*Z* =-8.97	.000
Puncture	6.57	2.00	124.46	8		
Recovery	2.49	2.89	75.53	1	*Z* =-6.42	.000

Table 4.11

Spearman correlations of pain scores assessed by two scales (n = 53)

	Clinical pain scale		
	Baseline phase	Puncture phase	Recovery phase
PIPP-R scale			
Baseline phase	.36*		
Puncture phase		.79*	
Recovery phase			.88*

Note. * $p < .001$.

4.1.3 Phase III Clinical Utility Evaluation

The frequency and percentage of each item are presented in Appendix G. Mean scores on each clinical utility dimension ranged from 3.66 to 3.93 (see Table 4.12). Most professional nurses rated "very good" to "good" level and the majority of nurses rated on "very good" for almost every item, except two. Those two items were related to requirement for training and not being certain about parent's satisfaction.

Table 4.12

Mean of each clinical utility dimension (n = 30)

Dimensions	Number of items	Mean
Appropriateness	5	3.66
Accessibility	2	3.93
Practicability	6	3.66
Acceptability	4	3.83

Summary

The clinical pain scale including four indicators was developed for pain assessment in clinical setting. The four indicators are heart rate changes, upper facial expression, lower facial expression, and sleep-wake states. The gestational age at birth was used to classify preterm neonates into two groups, less than 32 weeks and 32 weeks or over. The scales for these two groups have the same indicators and scoring, but the descriptions of behavioral indicators are different.

4.2 Discussion

This discussion is organized into three parts related to the study objectives including (1) characteristics of the newly developed clinical pain scale, (2) psychometric properties, and (3) clinical utility evaluation of the clinical pain scale for preterm neonates in the NICU.

4.2.1 Part I: Characteristics of the newly developed clinical pain scale

The results of this study demonstrate that the clinical pain scale can be used in the clinical setting to assess pain in preterm neonates. Characteristics of the clinical pain scale are described in terms of description in four pain indicators and scoring of the scale below:

Four pain indicators of the clinical pain scale

The clinical pain scale consists of four pain indicators can be categorized into two dimensions as behavioral indicators and physiological indicators. Behavioral indicators include upper facial expression, lower facial expressions, and sleep-wake states. The physiological indicator is an increase in heart rate. For better understanding, upper and lower facial expressions are described together and the rest of indicators are described separately below:

1) Upper and lower facial expressions

The unique feature of the clinical pain scale is the classification of upper and lower facial expressions into two age groups due to the fact that more mature infants display more vigorous facial expressions. In order to meet the objective of this scale in terms of clinical use, upper facial expression combined two facial actions including brow bulge and eye squeeze as one indicator. Lower facial expression combined two facial actions including nasolabial furrow and open mouth as one indicator. Even though in some literature called brow bulge, eye squeeze, nasolabial fold as upper facial response (Evans, 2001), nasolabial furrow was concerned as the lower facial expressions in the present study. Nevertheless, in neonates with mechanical ventilation, CPAP ventilation using nasal prongs, or cannula, most of facial actions may be influenced by the tubes or by the tapes used to fix the tubes, observation of upper and lower facial expressions was possible. The clinical pain scale separated facial expressions into two indicators, upper and lower facial expression, that differed from existing scales. Some existing scales such as NIPS scale and SUN scale used the whole face as one indicator leading to ambiguous and subjective determination. Some scales such as PIPP-R scale and BIIP scale used each facial expression leading to too many indicators to observe. Some other scales mixed both the whole face with each facial expression resulting in their complexity (e.g. COMFORTneo and COVERS). Separation of upper and lower facial expressions in this clinical pain scale made it easier for observation.

The I-CVI values of upper and lower facial expressions indicated that all experts agreed with using those indicators as a pain indicator for preterm neonates. In addition, the correlations of upper and lower facial expressions with total pain scores and between them were positive and high relationships indicated that they were good pain indicators for preterm neonates. Facial expression developed early in fetal life in the presence of any noxious stimulation and showed significant progress towards more complex facial expressions as gestational age increased (Reissland, Francis, & Mason, 2013). Facial expression of pain is a complex behavior and a spontaneous reflexive reaction to noxious stimuli, but it is, to a certain extent, subject to voluntary control (Chambers & Mogil, 2015). In later life as young as 8 years of age these reflexes become under the control of brain areas responsible for emotions. However, the preterm neonates are not capable to manipulate their facial expression of pain. The ability of an infant to display a change in facial expressions in response to a noxious stimulus requires that motor neuron activity is sufficiently coordinated to produce a visible set of facial muscle contractions (Slater et al., 2007). The facial nerve originates in pons and contains the special visceral efferents that innervate muscles of facial expressions. Brow bulge is caused by contraction of the frontalis and corrugator muscle. Eye squeeze is caused by contraction of the orbicularis oculi muscle. Nasolabial furrow is caused by contraction of the zygomaticus muscle. Immediately after painful stimuli and nociceptors activation, pain signals are sent through the spinomesencephalic tract and project to periaqueductal gray area at midbrain (Eiland, 2012). The periaqueductal gray area is extremely important in the modulation, interpretation, and response to pain. Stimulation of this area leads to changes in facial expression (Hall & Anand, 2005). Not only has the human study but also in the nonhuman animal species study proven quantifying pain through facial expression to have high accuracy and reliability (Chambers & Mogil, 2015). Thus, upper and lower facial expressions have been well established as a reliable and specific indicator of pain in preterm neonates.

2) Sleep-wake state

The unique feature of the clinical pain scale is the classification of sleep-wake state into two age groups due to the fact that more mature infants display more organized sleep-wake. Stable sleep-wake organization reflects the maturation of the CNS and state development involves increasing quiet sleep, decreasing active sleep, increasing awake, smooth and fewer transitions between wake and sleep, and increasing ability to sustain sleep periods with increased age (Foreman et al., 2008). Active sleep is the predominant behavioral state in the preterm neonate and is involved in maturation and brain development. In the fetal period, sleep is divided into stages, which are active sleep, quiet sleep, and indeterminate sleep. This is the first type of sleep existing in ontogeny, and it can already be identified in the week 28 to 30 of gestation. It is controlled by the forebrain and brainstem (Bonan, Pimentel Filho Jda, Tristão, Jesus, & Campos Junior, 2015). The mechanisms involved in sleep and wakefulness are present even before birth, and these are developed with circadian cycle, which is characterized by maintaining physiological functions. In the last trimester, the sleep-wake cycle of the fetus falls in line with maternal sleep-wake cycle, but the sleep-wake cycle of preterm neonates depends on the NICU environment. The interference of painful stimuli because of interconnections with the reticular activating system may play a role in the arousal effects of pain.

In addition to classification of sleep-wake state indicators into two age groups, the clinical pain scale used transition between sleep and wake states to differentiate behaviors of preterm neonates for three scoring levels. The scoring format and description made the scale easier for scoring than existing scales. Some existing scales such as SUN and COMFORTneo scales use five levels for CNS states or alertness scoring. The PIPP-R scale uses four-level scoring (active and awake, quiet and awake, active and asleep, and quiet and asleep) for behavioral states. However, the observed scored will be calculated for the total score of only if the sub-total scores of the physiological and facial indicators is more than zero; otherwise this behavioral states would be considered as a contextual factor. This scoring approach of the PIPP-R scale leads to its complexity.

The I-CVI values of sleep-wake states indicated that all experts agreed with using it as a pain indicator for preterm neonates. The sleep-wake state changes had positive correlation with both overall scores and facial expressions, which were the same dimensions. These correlations indicated that sleep-wake state changes was a good pain indicator. Even though preterm neonates with endotracheal tubes could not hear the cry to respond, the silent cry could be observed. The clinical pain scale included sleep-wake state changes as behavioral indicators instead of using behavioral state at baseline as factors influencing pain response like other scale (i.e., PIPP-R) because it reflected pain response of preterm neonates. Most infants were in an active sleep state and no neonates were crying during baseline phase. During the puncture phase, the infants showed high activity mainly crying, and stayed awake and alert until the dressing phase and significantly fewer reactions during recovery. The higher behavioral responses of sleep–wake state which were most apparent during the puncture phase were also found in previous studies (Chimello et al., 2009; Gaspardo, Chimellio, Cugler, Martinez, & Linhares, 2008; Walden et al., 2001).

3) An increased heart rate

An increased heart rate was only one physiological indicator that had been found useful for identifying pain response in the present study. In terms of distinguishing between pain and non-pain event and detecting pain level, three level scoring of heart rate was suitable for clinical setting. The result of the present study found that mean heart rate scores of the puncture phase (pain event) were higher than those of baseline and recovery phases (non-pain event) that was consistent with previous descriptive studies (Bozzette, 1993; Craig et al., 1993; Gibbins, Stevens, McGrath et al., 2008; Walden et al., 2001). The heart rate change from baseline at 30 seconds and 60 seconds immediately after the puncture in this study were 11.03 beats/minute and 15.97 beats/minute which closes to the previous study (Stevens et al., 1994; Xia et al., 2002). Theoretically, when the preterm infants are at rest and undisturbed, there are relatively few episodes of abnormally low heart rate or of abnormally high heart rate (Harrison, Roane, & Weaver, 2004), heart rate variability in the control period has never more than 5 beats/minute (McIntosh, van Veen, & Brameyer, 1994). Immediately after a painful stimuli and nociceptors activation, pain signals are sent through the spinorecticular tract and projected to the supraspinal area and trigger an autonomic response such as an increased heart rate (Eiland, 2012). Thus, heart rate change relative to heart rate at baseline is only one physiological indicator in the clinical pain scale.

Heart rate monitoring which is readily available in the clinical setting, and is an important parameter for the clinical management of critically ill neonates. The result of this study indicates that combing heart rate change indicator with the other indicators is specific enough to distinguish the scores between painful and non-painful events. The scoring of an increase heart rate of < 5, ≥ 5 to 9, and ≥ 10 beats/minute seems reasonable. In addition, using clearly defined term of “an increased heart rate” rather than “heart rate change” allows more understandable because heart rate change can be either a decreased heart rate or an increased heart rate.

The correlation coefficients between an increased heart rate indicator and three behavioral indicators were acceptable. Since an increased heart rate scoring was reviewed twice by content experts, the I-CVI value of it finally reached 1.00. From literature review related to pain pathway, an increase heart rate is autonomic response to the noxious stimuli. Nociceptive transmissions through A-delta and C-fiber synapse in the anterior horn (lamina VII) which stimulate preganglionic sympathetic neurons of the autonomic nervous system in the intermediolateral column to cause vasoconstriction (Evans, 2001). The ascending pain fiber connecting with reticular activation system and periaqueductal gray area at midbrain trigger protective autonomic response (increase in heart rate and respiratory rate) and facial responses. Taken together, results of this study showed evidence supporting that an increased heart rate was a good pain indicator for preterm neonates.

Three factors affecting pain reactivity, namely respiratory support, length of NICU stay, and previous pain exposure, were not included in the scale as expected. Respiratory support was eliminated based on suggestion of the experts. They recommended that it be collected as demographic information. With regard to length of NICU stay and previous pain exposure, they were excluded due to low values of inter-item correlation and item-total correlation. The plausible explanation of these low levels is that their scores before, during and after painful procedures remained the same. In addition, scoring of previous pain exposure had only two levels of score (0 or 1) that might lead to less variation of scores. Therefore, further study to reexamine these two factors is suggested.

Scoring of the clinical pain scale

Two concerning issues related to scoring in the clinical pain scale were the gestational age difference in behavioral responses and the delayed response of preterm neonates.

1) Gestational age difference in behavioral responses

With regard to difference in neuroanatooical maturity between preterm infants with GA < 32 weeks and those with GA ≥ 32 weeks, the developed clinical pain scale proposed different scoring of behavioral indicators for the two groups. Gestational age at birth was used as a criterion in this scale instead of postmenstrual age or postnatal age. The previous study found that the behavioral response in a painful event between two groups of infants born at 32 weeks' gestational age and at 28 weeks' gestational age was different. The preterm infants born at 28 weeks' GA showed less behavioral_responsiveness comparing with 32 weeks' gestational age (Johnston & Stevens, 1996). Therefore, scoring for behavioral response to pain of infant who has < 32 gestational age at birth and ≥ 32 gestational age at birth must be different.

The gestational age impacts on the ability of facial expression in preterm neonates. The preterm neonates aged > 32 weeks have the ability and maturity for a robust response (brow bulge & eye squeeze more than once, nasolabial furrow and open mouth widely, and movement with intense body and cry), whereas the younger might have less response. The comparison study of pain responses during the puncture phase in infants of different gestational age found that infants in the youngest gestational age strata (infants born less than $27^{6/7}$ weeks) had significantly less brow bulge ($p < .008$), eye squeeze ($p < .03$) and taut tongue ($p < .005$) than infants 32 to $35^{6/7}$ weeks' gestational age (Gibbins, Stevens, Beyene et al., 2008). The least mature infants had significantly less nasolabial furrow than all other gestational age strata ($p < .0001$) and less horizontal mouth stretch ($p < .001$) than infants 28 to $31^{6/7}$ weeks' gestational age and 32 to 35 weeks' gestational age.

The gestational age also impacts on the transition of sleep-wake states of preterm neonates. Become actively awake or at least brief arousal in young infants and fussing in older ones were common reactivity to pain in terms of sleep-wake states in preterm neonates (Gibbins, Stevens, McGrath et al., 2008; Grunau et al., 2001). In the present study, most infants from 53 observations were stable in the sleep or active sleep state in recovery phase. It can be explained that extremely low gestational age infants are physiologically immature and acutely ill, and as such may lack neurobehavioral organization to maintain awake states for long periods of time. As infants mature in gestation, their sleep–wake patterns become more regular and distinguishable (Gibbins, Stevens, McGrath et al., 2008). Most infants at 32 weeks' postconceptional age were in light sleep during baseline and shifted significantly from baseline to squeeze, with only 2% of infants in light sleep, 37% drowsy, and 55% highly aroused or agitated. The shift from squeeze to recovery was also significant, with 30% in light sleep and 51% drowsy (Grunau et al., 2001). In addition, the study by Foreman and colleagues (2008) recommended that gestational age of 32 weeks was a turning point for the maturity of distinct state behaviors because infant sleep-wake states were increasingly distinct after 32 weeks.

2) *Delayed response of preterm neonates*

The prominent point of scoring in the clinical pain scale is clearly instruction of scoring at 60 seconds after procedures or skin invasion. Concerning the delayed response of preterm neonates, this study expands the observation period to 60 seconds after puncture. The preterm neonates has delayed behavioral response and this delayed responses or absent response are always misinterpreted. An infant who undergoes painful stimuli and does not demonstrate a change in facial expressions cannot be assumed that the stimulus does not reach the cortex. Developmental changes underlying infant behavior are critically important if pain scores are to be correctly interpreted.

In a previous study using the PIPP scale to assess in 95 infants clinically required heel lances also found that 38% of 172 observations a lack of change in facial expressions (Slater et al., 2009). With the parallel use of two scales in this study, following the PIPP-R scale instruction by observing the infant for 30 seconds immediately following the event has shorter timing observation period than the clinical pain scale. There was 17.00% (9 of 53 occasions) lack of brow bulge and 20.80% (11 of 53 occasions) lack of eye squeeze, and 24.53% (13 of 53 occasions) lack of nasolabial furrow during the puncture phase assessed by PIPP-R scale. By scoring at 60 seconds after puncture assessed by the clinical pain scale, only 5.66% of observed occasions (5 of 53 occasions) lacked of upper facial expression and 13.21% of observed occasions (7 of 53 occasions) lacked of lower facial expression during puncture phase. It is clearly that the duration of observation is very important. It may be concluded that the clinical pain scale can better detect pain responses, especially facial expression compared with the previous scales and the previous studies.

The results of two previous studies using the near infrared spectroscopy (NIRS) in monitoring cortical activation (Bartocci et al., 2006; Slater, Cantarella et al., 2006) support the idea that scoring at 60 second after puncture can indicate pain in preterm neonates.

The first study recruited only 18 preterm neonates, some of them up to five times during development, with their postnatal ages ranging from 5 to 134 days and a period of 30 seconds following the heel lance (a peak response that occurred within 20 seconds) found increased selective contralateral cortical activation (Slater, Cantarella et al., 2006). In addition, another study by Slater and colleagues (2009) shows that the latency to facial response is longer and more variable in younger infants. There appears to be a marked stabilization of latencies at approximately 32 postmenstrual weeks. The second study recruited 40 preterm neonates (28 to 36 weeks) on the second day of life (25 to 42 hours of age, with each neonate being studied only once) and lasted between 35 seconds to 60 seconds following venipuncture, while a needle was being inserted into vein (average response that occurred over a 60-second-period) found that bilateral activation of the somatosensory cortex (Bartocci et al., 2006). The latency of pain response in preterm varies depending on their neurobiology of axonal and synaptic development. The lack of myelination and lower conduction velocities of C-fiber contribute to low speed of CNS processing in preterm infants. Thus the point of time for giving score should do after 20 to 60 seconds. Currently, Martakis and colleagues (2016) recruited 20 preterm neonates with gestational age less than 26 weeks (22^{+3} to 26 weeks) and observed their pain-related reactions on the third day of life with each neonate being studied only once. Fourteen infants presented a facial expression reaction. One infant with 23 weeks' gestational age showed nasolabial furrow as the first pain-related reaction with 10 seconds of latency time. One infant with 23 weeks and three days gestational age showed movements of forehead as the first pain-related reaction with 30 seconds of latency time. Therefore, scoring at 60 seconds after puncture can indicate pain in preterm neonates with gestational age of less than 26 weeks.

4.2.2 Part II: Psychometric testing

1) Reliability

The reliability testing for this study were internal consistency (homogeneity) and inter-rater reliability (equivalence).

1.1) Internal consistency

Cronbach's alpha coefficient of the clinical pain scale which consisted of four indicators was .86 for during procedure (n = 53) and .94 for all three phases (n = 159).

This values are acceptable for a new scale indicating that it had internal consistency. The reliability of the scale based on the degree of within scale indicator inter-correlation. Therefore, it indicated that those four indicators measuring the same thing and the same construct which was pain in preterm neonates.

1.2) Inter-rater reliability

The second aspect of reliability analysis was inter-rater reliability. The assessment of observer agreement was assessed with the ICC (3,k). The value of ICCs on total scores obtained before, during, and after painful procedures yielded correlations were .98, .91, and 1.00, respectively indicating that the new clinical pain scale met inter-rater reliability. There is no standard acceptable level of reliability using the ICC. It ranges from zero (implying no agreement) to one (perfect agreement), with value close to one representing the higher reliability. Chinn (1991) recommends that ICC of at least .6 to be useful. The results of the present study indicated the equivalence of the clinical pain scale when different raters used the scale to assess pain in preterm neonates. Among three phases, there was a good agreement between two raters of pain scores assessed by the clinical pain scale. The clear instruction of the time point for scoring and combining facial expressions leading to gain a reasonable number of indicators might be a reason for this reproducibility and repeatability.

2) Evidence of validity

Construct validity is the degree to which an instrument measures the construct it is intended to measure. The present study measured construct validity in two approaches including hypothesis testing and concurrent validity.

2.1) Evidence of construct validity: hypothesis testing

Construct validity was established by comparing the pain scores on the clinical pain scale during a painful event (puncture phase) and non-painful event (baseline and recovery phases). The results of this study demonstrated a significant difference in pair scores between painful and non-painful event ($p < .001$) indicating that the scale can discriminate between levels of pain across three phases of painful procedures. The score differences were in the expected direction of the hypotheses testing, thus construct validity is supported. Due to the theoretical meaning of the measurements based on the concept of pain in preterm neonates which was derived from the process of scientific inquiry, the concept analysis. Moreover, indicators of the clinical pain scale were generated by clinical observations and interviews. The results of items analysis affects both reliability and validity by manipulating the variability of scores, eliminating the extraneous effects of non-relevant indictors, and strengthening the relationship between indicators (Jacobson, 1997). Beside proper indicators for pain in preterm neonates, calibration of monitor was done regularly as the standard checked up for controlled systematic error. Two ways to minimize measurement error with physical instruments are to verify well calibrated equipment and to maximize reliability (Houser, 2008)

A low pain score (not equal to zero) during non-pain events (baseline and recovery phases) can be explained by two reasons. First, handling for antisepsis prior to puncture may cause increase in behavioral and physiological responses. Second, the early and repeated exposure to painful experiences reduce infants' pain threshold and provoke hyperalgesia and allodynia. Consequently, the neonate would exhibit behavioral and physiological responses during painless procedures that are similar to those during painful ones. However, the higher behavioral responses (facial activities and sleep–wake state) were most apparent during the puncture phase.

2.2) Evidence of construct validity: convergence validity

Convergence validity is defined as the extent to which a test yields the same results as other measures of the same phenomenon. Convergence validity was established by comparing the clinical pain scale to previously validated pain scales, namely the PIPP-R, and demonstrating a positive and high degree of correlation across three phases of painful procedures. This correlation supports the evidence of convergence validity of the scale. Because the idea of convergence is that different measures of the same trait should correlate highly with other one. Validity correlations are usually lower than those for reliability; relationship of r equal to .40 to .60 may be entirely satisfy (Jacobson, 1997).

4.2.3 Part III: Clinical utility evaluation

Clinical utility in this study was evaluated based on the users' judgment on its appropriateness, accessibility, practicability, and acceptability. Regarding to those four aspects, the impact of the clinical pain scale for preterm neonates on pain management, availability, suitability for using in NICU environment, and the preference of users was examined. It has already been shown that the clinical pain scale was well accepted and in fact preferred by NICU nurses, except issue related to training. According to the negative question, “a user can use a scale without training”, most nurses disagreed with it and indicated they needed to be trained before using a scale. The bedside assessment duration took 90 seconds. But, exact time of scoring were not monitored. The group training duration took approximately one hour. Training session was designed by using case scenario videos for demonstrating scoring, practicing especially in facial expression observation, and testing. The researcher provided communication channel such as a free communication application and telephone number for asking question in case they had any problem with using the scale. During data collection period, all trained nurses had no question during assessment. It might be concluded that they clearly understood how to use the scale after one-hour training. Due to the multi-dimension of the scale, there are many things that need to be observed as the same time. Anand (2007) stated that the behavioral and physiological responses required the subjective evaluation of a clinical observer. Therefore, NICU nurses need time for practicing and training how to observe and score all indicators, especially facial expressions.