



**Psychometric Properties of Pain Intensity Scales Comparing Among
Postoperative Adult Patients, Elderly Patients Without
and With Mild Cognitive Impairment in China**

Yinghua Zhou

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Nursing Science (International Program)**

Prince of Songkla University

2010

Copyright of Prince of Songkla University

Thesis Title Psychometric Properties of Pain Intensity Scales Comparing Among Postoperative Adult Patients, Elderly Patients Without and With Mild Cognitive Impairment in China

Author Miss Yinghua Zhou

Major Program Nursing Science (International Program)

Major Advisor:

.....
(Assist. Prof. Dr. Wongchan Petpichetchian)

Examining Committee:

..... Chairperson
(Assoc. Prof. Dr. Praneed Songwathana)

Co-advisor:

.....
(Dr. Luppana Kitrungrote)

.....
(Assist. Prof. Dr. Wongchan Petpichetchian)

.....
(Dr. Luppana Kitrungrote)

.....
(Assist. Prof. Dr. Sasikaan Nimmaanrat)

.....
(Dr. Supaporn Wannasuntad)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Nursing Science (International Program)

.....
(Assoc. Prof. Dr. Kerkchai Thongnoo)
Dean of Graduate School

Thesis Title Psychometric Properties of Pain Intensity Scales Comparing
Among Postoperative Adult Patients, Elderly Patients
Without and With Mild Cognitive Impairment in China

Author Miss Yinghua Zhou

Major Program Nursing Science (International Program)

Academic Year 2009

ABSTRACT

In China, there is a lack of studies to generate pain intensity scales to the patients with CI and compare the pain scales in various age groups. Therefore, this study would compare the psychometric properties of the evidence-supported pain intensity scales consisting of the Verbal Descriptor Scale (VDS), the Numeric Rating Scale (NRS), the Faces Pain Scale (FPS), the Numeric Box-21 Scale (BS-21), and the Colored Analogue Scale (CAS) in Chinese postoperative adults varying in ages including the elderly with mild CI. This was a descriptive comparative study and 200 surgical patients were recruited purposively from a university-affiliated hospital with 50 for each group: young adults (age 20 – 44 years), middle-aged adults (age 45 – 59 years), elderly (age \geq 60 years) without CI, and elderly (age \geq 60 years) with mild CI. Participants rated the vividly remembered, current, worst, least, and average pain, and indicated scale preference and simplicity. Scale face validity, concurrent validity, convergent validity, and test-retest reliability at a 3-day interval were assessed. Fisher's exact tests were used to investigate whether face validity was related to different age

groups and the levels of CI. One-way ANOVA and Kruskal-Wallis test were used to test the differences of concurrent validity, convergent validity, and test-retest reliability of each pain scale among the four groups. Regarding face validity, the FPS was ranked best across the subjects as nearly half of the patients selected it as both the most preferred and simplest and it had low errors; the VDS and the NRS were similar and ranked following the FPS; however, the BS-21 and the CAS were ranked last. The concurrent validity, convergent validity, and test-retest reliability of all five pain scales were supported in use with the four groups. The differences in psychometric properties among the four groups were only found in face validity. The findings support the psychometric properties of all five pain scales for pain assessment in Chinese adults including the elderly with mild CI. However, the FPS appears to be the best scale followed by the VDS and the NRS.

ACKNOWLEDGEMENTS

It is impossible to finish this work without the continuous help and love from the people around me. I take this opportunity to give thanks to them for their willing to teach, guidance, help, and care during my master degree study.

I begin by expressing my great gratitude to my wonderful major advisor, Assist. Prof. Dr. Wongchan Petpichetchian who has continuously guided me, sharpened my thinking, and cared me so much during these two years. Her expertise in research methodology and statistical analysis and love teaching principle give me invaluable experience for my life. The second person that I want to give thanks is my co-advisor Dr. Luppana Kitrungrote. Her kindness, patience, and guidance make the whole experience joyful for me. I am so lucky to be their advisee. I would like to thank all the examining committee members for their valuable suggestions. I also wish to thank all the teachers in Faculty of Nursing, Prince of Songkla University for their willing to teach and guidance during my study.

I also want to express my appreciation to all the people who contribute to my study including all the participants, nurses, and doctors in the Second Affiliated Hospital of Kunming Medical University. Special thanks to my head nurse, Qing Zhang, and nurse director, Huiyun Yuan, for their great assistance in the data collection process. Special appreciation to all the experts in evaluation and validation of the instruments, Assoc. Prof. Dr. Praneed Songwathana, Assoc. Prof. Dr. Sasitorn

Phumdoung, and Assist. Prof. Dr. Sasikaan Nimmaanrat. I also wish to thank Mr. Withoon Sangkharak for helping me to design the picture of my instrument.

I wish to thank Prince of Songkla University for awarding me the MOU scholarship and the teaching assistant scholarship which give me strong financial support and invaluable teaching experience for my career. I also wish to thank all the staffs in Faculty of Nursing, Prince of Songkla University. They give me continuous guidance and help during these two years study.

I also want to give thanks to all my study colleagues and friends. Special thanks to Pannee, Duangsuda, and Nattaya and my Chinese friends Yun Zhang and Yan Zhuang. I also want to thank the PhD students for giving me help and advice.

Lastly but not least, I wish to express my great gratitude to my family and relatives for their endless love, understanding, and strong support in every step of my life. Since I write this acknowledge on the 2010 Chinese Spring Festival, I want to give my wishes to my family and relatives, teachers, friends, and participants. I wish that they keep healthy and have a happy life!

Yinghua Zhou

CONTENTS

	PAGE
CONTENTS	vii
LIST OF TABLES	xi
LIST OF FIGURES	xiii
CHAPTER	
1. INTRODUCTION.....	1
Background and Significance of the Study.....	1
Objectives of the Study.....	6
Research Questions of the Study.....	8
Hypotheses of the Study.....	9
Theoretical Framework of the Study.....	10
Definition of Terms.....	12
Significance of the Study.....	18
2. LITERATURE REVIEW.....	19
Pain in Postoperative Patients.....	19
Concept of Pain.....	19
Pain Mechanisms	21
Impacts of Unrelieved Postoperative Pain	23
Postoperative Pain Management.....	24

CONTENTS (continued)

	PAGE
Factors related to Pain experience.....	27
Measurement and Pain Assessment Measures.....	33
Definition of Measurement and Measurement Theory.....	33
Psychometric Properties of Instrument.....	34
Pain Assessment Tools.....	56
Factors Influencing Psychometric Properties of Pain Scales.....	71
Cognitive Impairment.....	77
Concept.....	77
Cognitive Impairment, Tool Selection, and Psychometric Properties of Pain Scales.....	78
Psychometric Properties of Pain Intensity Scales in Elderly With Cognitive Impairment	81
Cognitive Function Assessment.....	82
Treatment for the Mild Cognitive Impairment.....	88
Impacts of Surgery on Cognitive Function	89
Summary.....	90
3. RESEARCH METHODOLOGY.....	93
Research Design.....	93

CONTENTS (continued)

	PAGE
Research Setting.....	93
Population and Sample	96
Inclusion Criteria.....	97
Instrumentation.....	98
Instruments for Collecting Background Data.....	99
Instruments for Assessing Patients' Pain Intensity.....	100
Instruments for Assessing the Validity of the Pain Intensity Scales.....	101
Validity and Reliability of the Instruments.....	103
Ethical Considerations.....	104
Data Collection Procedures.....	105
Data Analysis.....	109
4. RESULTS AND DISCUSSIONS	111
Results.....	111
Discussions	123
5. CONCLUSIONS AND RECOMMENDATIONS	132
Summary of the Study Findings.....	132
Strengths and Limitations	133
Implications and Recommendations	134

CONTENTS (continued)

	PAGE
REFERENCES.....	137
APPENDICES.....	152
A: Informed Consent Form.....	153
B: Instruments.....	156
B1: Demographic Data Questionnaire.....	157
B2: Chinese version of Mini-Mental State Examination.....	159
B3: Patient’s Pain Intensity Assessment Form One.....	161
B4: Patient’s Pain Intensity Assessment Form Two.....	163
B5: Modified Pain Interference Scale	169
B6: Scale Preference Questionnaire.....	171
B7: Scale Simplicity Questionnaire.....	173
B8: Scale Accuracy Checklist.....	175
C: Tables.....	176
D: Methods in Data Analysis	178
E: List of Expert Participants	179
VITAE	180

LIST OF TABLES

Tables	PAGE
1. Frequency and Percentage of Demographic Characteristics	112
2. Preference of the Five Pain Intensity Scales in All Subjects	115
3. Test of the Most Preferred Scales Classified by Different Age Groups Levels of CI, Using Fisher's Exact Test	115
4. Simplicity of the Five Pain Intensity Scales in All Subjects	116
5. Test of the Simplest Scales Classified by Different Age Groups and Levels of CI, Using Fisher's Exact Test	116
6. Accuracy of Five Pain Intensity Scales in All Subjects	117
7. Test of Scale Accuracy Classified by Different Age Groups and Levels of CI, Using Fisher's Exact Test	117
8. Concurrent Validity: Pearson Product-moment Correlation Coefficients Between the Scores of the Pain Interference Scale and Scores of Each of the Five Pain Intensity Scales in All Subjects	119
9. Test of the Difference of Concurrent Validity Comparing Among Different Age Groups and Levels of CI, Using Kruskal-Wallis Test	119
10. Convergent Validity: the Pearson Product-moment Correlation Coefficients Between the Scores of the VDS and the Scores of	

LIST OF TABLES (continued)

Tables	PAGE
Each of the Four Pain Intensity Scales in All Subjects	121
11. Test of the Difference of Convergent Validity Comparing Among Different Age Groups and Levels of CI, Using Kruskal-Wallis Test	121
12. Test-retest Reliability: Correlation Coefficients Between the Vividly Remembered Pain Ratings at a 3-day Interval in All Subjects	122
13. Test of the Difference of Test-retest Reliability Comparing Among Different Age Groups and Levels of CI, Using One-way ANOVA	123

LIST OF FIGURES

Figures	PAGE
1. Theoretical Framework of This Study	12
2. Data Collection Procedure	108

CHAPTER 1

INTRODUCTION

Background and Significance of the Study

Pain is a common problem which is especially true in postoperative patients (Nendick, 2000). Despite progress in pain management, some studies show that postoperative pain is still undertreated (Apfelbaum, Chen, Mehta, & Gan, 2003; Karanikolas & Swarm, 2000). Research findings also demonstrated a high prevalence of unrelieved pain in Chinese patients undergoing surgery (Chung & Lui, 2003; Shen, Sherwood, McNeill, & Li, 2008). For example, Shen et al. (2008) investigated the outcome of postoperative pain management from the second-day-postoperative inpatients and found that 78% of the patients reported pain in the past 24 hours and the mean ratings for pain were moderate to severe.

In addition, postoperative pain management is so important that we need to pay attention to, especially in elderly. In China, the percentage of populations aged over 60 years is 10.4% in 2000 and is expected to reach 19.7% in 2025 (Zhang, Cheng, & Chen, 2001). As the population continues to age, the number of elderly surgical patients will increase. Moreover, the physical function in elderly is declined and the elderly have an increased vulnerability to stressors (Aubrun & Marmion, 2007). It is increasingly clear that elderly are at greater risk for pain-related morbidity including physical disability and psychological distress, such as myocardial infarction,

respiratory failure, depression, and delirium (Aubrun & Marmion; Gagliese & Melzack, 2006).

It is well established that reliable and valid pain assessment is a prerequisite for effective pain management (Sjostrom, Dahlgren, & Haljamae, 1999). The assessment of pain is complex as pain is a multidimensional experience; however, pain intensity is probably the most frequently assessed component of pain (Gagliese & Melzack, 2006) because pain intensity is used to determine the effectiveness of pain management in clinic. In addition, self-report is the most reliable way for accurate assessment of pain due to its subjectivity (Pasero, 2003). Therefore, pain intensity scales that based on self-report are the commonly used tools for assessing pain. In clinic, there are many pain intensity scales to quantify pain assessment and guide treatment.

The most commonly used pain intensity scales are the Visual Analog Scale (VAS), the Numeric Rating Scale (NRS), and the Verbal Descriptor Scale (VDS) in western countries. Some studies show that the VDS and the NRS may be appropriate across the adult lifespan; for the VAS, though it has high validity, there are many limitations especially when used in elderly patients including high rates of unscorable data and low face validity (Gagliese, Weizblit, Ellis, & Chan, 2005; Herr, Spratt, Mobily, & Richardson, 2004; Peter, Patijn, & Lame, 2007). Besides the commonly used three scales, many studies find that the Faces Pain Scale (FPS) and the Numerical Box-21 Scale (BS-21) are also suitable for pain assessment in younger and

older adults. Herr, Mobily, Kohout, and Wagenaar (1998) supported the validity and reliability of the FPS when they evaluated the FPS in elderly. Another study by Peters et al. (2007) compared the VAS-horizontal (VAS-H), the VAS-vertical (VAS-V), the VDS, the Numerical Box-11 Scale (BS-11), and the BS-21 in younger and older patients. They found that the BS-21 was highly valid, had the lowest number of mistakes, and was the most preferred scale overall, although patients aged over 75 years especially preferred the VDS. Therefore, they recommended that the BS-21 is the first choice for pain intensity assessment in heterogeneous patient groups and the VDS can be considered when the majority of patients are older adults. In sum, the NRS, the VDS, the FPS, and the BS-21 might be suitable for assessing pain in younger and older adults; however, the VAS might be problematic especially when used in elderly.

Pain assessment in elderly with cognitive impairment (CI) presents challenges to health care providers because of the communication disorders, the characteristics of available pain measures, and the reduced tendency to report pain in these patients (Stolee et al., 2005). As CI deteriorates, patients are less capable to use the self-report pain intensity scales and pain behavioral observation tools are more likely to be preserved (Hadjistavropoulos & Craig, 2002). However, some studies show that patients with mild or moderate CI can use pain intensity scales to report pain reliably and validly but the qualities of pain scales are varied. Closs et al. (2004) compared the VDS, the NRS, the FPS, the Colored Analogue Scale (CAS), and the Mechanical Visual Analog Scale (MVAS) among elderly with different levels of CI.

They found that the VDS was the most easily understandable scale and appeared to be suitable for all but the elderly with severe CI; the NRS was the second most successful scale following the VDS. In contrast, Chibnall and Tait (2001) compared the VDS, the FPS, the horizontal 21-point (0-100) box scale (BS-21), and the vertical 21-point (0-20) box scale in elderly without and with CI. They found that the BS-21 emerged as the best scale regarding psychometric properties and supported the use of the BS-21 for pain assessment in elderly including those with mild to moderate CI. However, when Taylor and Herr (2003) compared the VDS, the NRS, the FPS, and the Iowa Pain Thermometer (IPT) in African American elderly without and with CI, they found that all scales were easy to use, valid, and reliable in both elderly without and with CI groups but both groups preferred the FPS to report their pain intensity. The CAS is another tool that might be feasible for using in elderly with CI. Scherder and Bouma (2000) compared the CAS, the FPS, and the Facial Affective Scale (FAS) among an early and midstage of Alzheimer's disease (AD) population and nondemented elderly population. They found that the CAS worked best and could be comprehended very well in both nondemented elderly patients and the early AD patients. In sum, based on the above evidences, the VDS, the NRS, the FPS, the BS-21, and the CAS can be attempted to assess pain in the patients with none to mild or moderate CI.

Cultural background, one aspect of the sociocultural dimension of pain, has been identified as an important element that influences pain behavior and expression (Lasch, 2000; McGuire, 1992). Moreover, several studies show that culture

also influences the psychometric properties of pain assessment scales. One study (Herr et al., 1998) evaluated the construct validity of the FPS for use with the European-American elderly. They found that the subjects agreed that the FPS represented pain more strongly than any of other constructs; however, the faces could also represent other constructs except anger. In addition, only significant difference between the construct of anger and all other concepts (e.g., pain, sourness, sleepiness, sadness, and boredom) was found. However, when Taylor and Herr (2002) used the FPS with African-American elderly, they found that the construct of pain was significantly different from the constructs of sourness, sleepiness, and boredom, but not different from sadness and anger. Therefore, cultural background can influence the construct validity of the FPS. Moreover, in China, one study found that Chinese patients assessed pain more accurately with a vertical version of the VAS than with the more commonly used horizontal version (Aun, Lam, & Collett, 1986).

Considering the cultural difference, however, studies on pain scale use are limited in China. One study, investigating the correlation between the NRS and the VDS in 50 Chinese postoperative patients, found that the NRS and the VDS correlated well and suggested that a scale incorporating an NRS with a VDS was recommended (Zhao, Lu, Zhao, & Tou, 2002). Another study by Li, Liu, and Herr (2007) compared the VAS, the NRS, the VDS, and the Faces Pain Scale Revised (FPS-R) in Chinese postoperative adults. They found that all four scales had good validity and reliability, both the VDS and the FPS-R had low error rates, and nearly half of patients preferred

the FPS-R, but the VAS was difficult to understand and had highest error rates. Finally, they recommended that the FPS-R appeared to be the best scale for Chinese adults. However, in this study, the evidence-supported pain scales BS-21 and CAS were not included. In addition, the cognitive function was not measured in this study and the patients who could not complete the pain scales were excluded during the data analysis.

In conclusion, the findings of western studies showed that the qualities of pain assessment scales use in various age groups including elderly with CI are varied. In addition, evidences from western countries support the VDS, the NRS, the FPS, the CAS, and the BS-21 as options for assessing pain in patients including those with none to mild or moderate CI. However, in China, there is a lack of studies to generalize these scales to patients with CI. Therefore, considering the lack of studies with Chinese population to guide the decision which pain scale might work best, this study would compare the psychometric properties of the evidence-supported tools: the VDS, the NRS, the FPS, the CAS, and the BS-21 in the adults varying in ages including the elderly with mild CI in Chinese population.

Objectives of the Study

1. To determine the face validity of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and

elderly patients with mild CI.

2. To determine the concurrent validity of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

3. To determine the convergent validity of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

4. To determine the test-retest reliability of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

5. To compare the differences of validity and reliability coefficients of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

5.1 To compare the differences of validity coefficients of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

5.2 To compare the differences of reliability coefficients of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

Research Questions of the Study

1. What are the levels of face validity of each of the pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI?

2. What are the levels of concurrent validity of each of the pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI?

3. What are the levels of convergent validity of each of the pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI?

4. What are the levels of test-retest reliability of each of the pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without

CI, and elderly patients with mild CI?

5. What pain intensity scale, including the VDS, the NRS, the FPS, the CAS, and the BS-21, does offer the best validity and reliability for postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI?

5.1 What pain intensity scale, including the VDS, the NRS, the FPS, the CAS, and the BS-21, does offer the best validity for postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI?

5.2 What pain intensity scale, including the VDS, the NRS, the FPS, the CAS, and the BS-21, does offer the best reliability for postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI?

Hypotheses of the Study

1. There are differences in the validity coefficients of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

2. There are differences in the reliability coefficients of the five pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among

postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI.

Theoretical Framework of the Study

Self-report is the most reliable way for accurate assessment of pain due to its subjectivity (Pasero, 2003). Although pain is a multidimensional experience, pain intensity is probably the most frequently assessed component of pain (Dahl, 1996; Gagliese & Melzack, 2006) since pain intensity is used to determine the effectiveness of pain management in clinic. Therefore, this study would assess five evidence-supported pain intensity scales that based on self-report including the VDS, the NRS, the CAS, the FPS, and the BS-21.

In order to quantify a subjective construct, such as pain, the psychometric properties of its instrument need to be assessed. Validity and reliability are two major criteria for assessing the psychometric properties of an instrument. In this study, validity would be evaluated by face validity, concurrent validity, and convergent validity and reliability would be evaluated by test-retest method. Face validity refers to whether the instrument looks as though it is measuring the appropriate construct. Concurrent validity is one type of criterion-related validity which concerns using an instrument to estimate criterion behavior that is external to the measuring instrument itself. Concurrent validity is assessed by correlating a measure and the criterion within a short period of time when the criterion exists in the

present. Convergent validity refers to the correlation between different methods that measure the same trait. Test-retest reliability is one way to estimate the reliability of empirical measurements in which the same test is given to the same people after a period of time.

There are many factors that influence pain measurement. However, among those factors, age and cognitive level would be emphasized in this study. The reasons are as follows. For one thing, the findings of age differences in the psychometric properties of the pain scales are inconsistent. Some studies found that age differences in the psychometric properties of the pain scales were evident (Gagliese & Katz, 2003; Peters et al., 2007). However, another study failed to find the age differences in the psychometric properties of the pain scales (Li et al., 2007). Moreover, the findings of western studies showed that the qualities of pain scales use in various age groups including the elderly with CI are varied. Therefore, age factor was a focus in this study to determine whether age influenced the psychometric properties of the pain scales. For another, cognitive dimension is one dimension of pain. Cognitive level influences psychometric properties of pain assessment scales and the selection of pain assessment scales in clinical settings (Defrin, Lotan, & Pick, 2006; Scherder & Bouma, 2000). Pain in elderly with severe CI may be assessed by using behavioral observation tool which has been used as a mean of quantifying pain and discomfort (Hurley, Volicer, Hanrahan, Houde, & Volicer, 1992; Simons & Malabar, 1995). However, patients with mild CI may not get much benefit because of the

complex of the behavioral observation scale and time costing by the health care providers. Therefore, in this study, cognitive level was another focus to determine whether the pain scales were valid across the population of postoperative adult patients, elderly patients without CI, and elderly patients with mild CI.

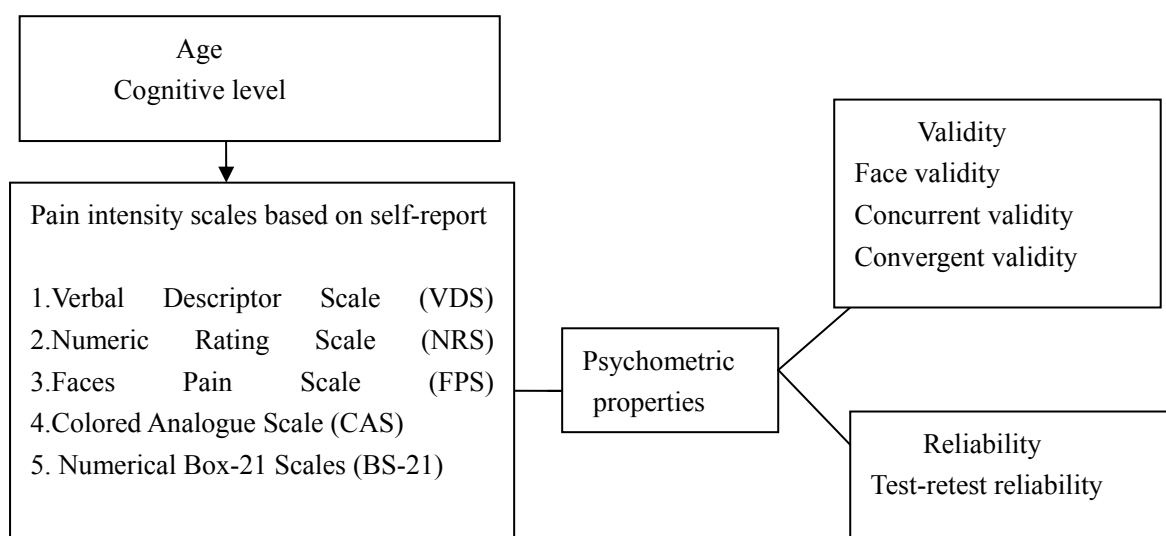


Figure 1. Theoretical framework of this study

Definition of Terms

Pain Intensity Scales

The pain intensity scales in this study consisted of the Numeric Rating Scale (NRS), the Verbal Descriptor Scale (VDS), the Colored Analogue Scale (CAS), the Numerical Box-21 Scale (BS-21), and the Faces Pain Scale (FPS).

Numeric Rating Scale (NRS)

The NRS is a line marked with 11 numbers (0 through 10) at equal interval where 0 is “*no pain*” and 10 is “*worst pain*”. Participants were asked to choose the number that best reflected their pain intensity.

Verbal Descriptor Scale (VDS)

The VDS used in this study consists of five adjectives that describe different levels of pain intensity: no pain, slight pain, moderate pain, severe pain, and unbearable pain (Pesonen, Suojaranta-Ylinen, Tarkkila, & Rosenberg, 2008). Participants selected the word that best represented their pain intensity.

Colored Analogue Scale (CAS)

The CAS is designed to assess pain intensity among children (McGrath et al., 1996). In this study, the CAS was modified in order to make it practical to be presented on the questionnaire. The modified CAS consisted of a vertical triangular shape, varying in width and hue and ranging from 1 cm wide and light yellow hue at the bottom to 2.5 cm wide and deep red hue at the top. These colors were selected based on the patients’ preference from a pilot study. The length of the triangular shape is 10 cm with anchors “0” and “*no pain*” at the bottom and anchors “10” and “*worst pain*” at the top. The participants were asked to mark a horizontal line on the triangular shape where best reflected their pain intensity and pain intensity was scored by

measuring the vertical distance from the bottom to the patient's mark.

Numerical Box-21 Scale (BS-21)

The BS-21 has a horizontal row of 21 boxes with numbers labeled from 0 to 100 in increments of five (Jensen, Miller, & Fisher, 1998). There are anchors “*no pain*” on the left extreme and “*worst pain*” on the right extreme. The participants selected the box that represented their pain intensity.

Faces Pain Scale (FPS)

The FPS (Bieri, Reeve, Champion, Addicoat, & Ziegler, 1990) consists of seven line-drawn faces presented in a horizontal format with different facial expressions that include a neutral face representing no pain, a severely controlled face without tears representing worst pain, and five other facial expressions in between. Participants selected the face that best reflected their pain intensity.

Psychometric Properties of the Instruments

Psychometric properties of the instruments are defined as validity and reliability. In this study, validity was evaluated by face validity, concurrent validity, and convergent validity and reliability was evaluated by test-retest reliability.

Face Validity

Face validity refers to whether the instrument looks as though it is measuring the appropriate construct. In this study, face validity of the five pain intensity scales was assessed by three aspects including preference, simplicity, and accuracy and these three aspects were assessed separately. Firstly, preference was assessed by using the Scale Preference Questionnaire (SPQ). On the SPQ, five pain intensity scales were presented and the participants were asked to rank order from 1 to 5 when 1 = most preferred and 5 = least preferred. Secondly, simplicity was assessed by using the Scale Simplicity Questionnaire (SSQ). On the SSQ, five pain intensity scales were presented and the participants were asked to rank order from 1 to 5 when 1 = simplest and 5 = least simple. Lastly, accuracy was evaluated by the number of subjects with accurate response (without any errors) and inaccurate responses (with any errors) by the researcher using the Scale Accuracy Checklist (SAC). Thus, in this study, the higher number of subjects indicating the most preferred scale, the simplest scale, and the higher number of subjects with accurate response of the scale indicated the higher face validity.

Concurrent Validity

Concurrent validity is one type of criterion-related validity which concerns using an instrument to estimate criterion behavior that is external to the measuring instrument itself. Concurrent validity is assessed by correlating a measure

and the criterion within a short period of time when the criterion exists in the present. In this study, since postoperative pain would interfere with patients' functioning, the modified Pain Interference Scale (PIS) was used as the criterion to assess the concurrent validity of each pain intensity scale. Concurrent validity was assessed by calculating Pearson product-moment correlation coefficient between the scores of the modified PIS and the scores of each pain scale by using the recalled worst pain during the past 24 hours. The higher is the correlation, the higher the concurrent validity.

Convergent Validity

Convergent validity refers to the correlation between different methods that measure the same trait. The recalled worst pain score of the VDS was used as the gold standard score to be correlated with the recalled worst pain scores on the NRS, the FPS, the CAS, and the BS-21 in this study since the VDS was thought to be easier for elderly to understand and the prior psychometric validation as mentioned in chapter two. Convergent validity was assessed by calculating Pearson product-moment correlation coefficient between the VDS and each of the four pain scales by using the recalled worst pain during the past 24 hours. The NRS is another tool that is easily to understand and is recommended for pain assessment in elderly including those with none to mild or moderate CI from a literature review. Therefore, if the correlation between the NRS and the VDS is high enough it will provide evidence for convergent validity of both the NRS and the VDS. For other three scales, the higher is the

correlation, the higher the convergent validity of that scale.

Test-retest Reliability

Test-retest reliability is one way to estimate the reliability of empirical measurements in which the same test is given to the same people after a period of time. Test-retest reliability of the pain scales was examined by the way that subject was asked to rate the intensity of the vividly remembered painful experience ever felt in his/her life by using the five pain intensity scales on the preoperative day and the 3rd postoperative day. Test-retest reliability was assessed by calculating Pearson product-moment correlation coefficient and Spearman rank correlation coefficient between the scores that obtained twice. The higher is the correlation, the higher the test-retest reliability.

Age Group

The subjects in this study were classified into three age groups: young adults (age 20 - 44 years), middle-aged adults (age 45 - 59 years), and old adults (age \geq 60 years) (Li, Bai, Liu, & Li, 2000; Yin, Sun, Yang, Hu, & Chen, 2000).

Mild Cognitive Impairment

Mild cognitive impairment (CI) is known as a transitional stage between normal ageing and early dementia. The Chinese version of Mini-Mental State

Examination (CMMSE) (Zhang, 1998) was used to assess the cognitive status of the patients in this study. The cutoff points of the CMMSE from the prior studies (Yang et al., 2008; Zhang et al., 1999) were used to divide the elderly into elderly without CI group and elderly with mild CI group. The elderly without CI group included illiterate people with score > 19, people having primary school education with score > 22, and people having secondary school education or higher with score > 26. The elderly with mild CI group included illiterate people with score 17-19, people having primary school education with score 20-22, and people having secondary school education or higher with score 24-26.

Significance of the Study

This study provides invaluable evidence to help nurses and other health professionals to select the best pain assessment tool, resulting in valid and reliable information about pain intensity. Consequently, patients can receive better pain management.

CHAPTER 2

LITERATURE REVIEW

The literature review started with the major concepts of this study: pain in postoperative patients, measurement, and pain assessment measures. These were followed by discussion the factors influencing pain measurement. In the above factors, age and cognitive impairment which were two focuses in this study were discussed. In addition, the theoretical and methodological considerations were also addressed.

Pain in Postoperative Patients

Concept of Pain

Pain is a common phenomenon in nursing practice but it is also a very complex concept to define. McCaffery (1981) defined pain as “whatever the experiencing person says it is and existing whenever the person says it does” (p.7). This definition emphasizes that pain is a subjective experience in nature and all pain is real even if its cause is unknown. The International Association for the Study of Pain (IASP) subcommittee on Taxonomy (1986) defined pain as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage”(International Association for the Study of Pain, 1994). This definition indicates that not only physical factors but also sensory,

affective, and motivational factors influence pain experience (McGuire, Kim, & Lang, 2004).

Ahles, Blanchard, and Ruckdeschel (1983) described pain as a multidimensional experience consisting of five aspects which were physiological, sensory, behavioral, cognitive, and affective. McGuire (1992) added a sociocultural dimension to this concept and described the components for each dimension. The physiological dimension includes the etiology, duration, location, endogenous opioids and neurotransmitters, and psychophysiologic factors; the sensory dimension includes the pattern, intensity, and quality of pain; the affective dimension includes emotional response (e.g., depression, anxiety, worry, helplessness, mood, fear), suffering, and psychiatric disorders; the cognitive dimension includes the view of self, meaning of pain, coping strategies, attitudes, beliefs, knowledge, influencing factors, level of cognition, and pain relief; the behavioral dimension includes indicators of pain, pain control behaviors, communication of pain, and related symptoms such as fatigue and sleep; the sociocultural dimension includes demographic variables, cultural background, personal, family, work roles, family factors, and caregiver perspectives.

In conclusion, the clinicians should believe that pain is a subjective feeling that only can be perceived directly by the sufferers. Pain is also a multidimensional experience that can be influenced by physiological, sensory, behavioral, cognitive, affective, and sociocultural factors.

Pain Mechanisms

Nociception

Nociception is the physiological theory which is derived from other disciplines and is applied in pain area by nurse researchers. Nociception is the mechanisms that encode and transmit the pain signal, via a complex series of ascending pathways, from the peripheral nerve endings (nociceptors) to the brain (Renn & Dorsey, 2005). In the cerebral cortex, the nociceptive signals (pain signals) are interpreted as pain.

Gate Control Theory

The gate control theory was developed by Melzack and Wall in 1965 and many studies continue using this theory for understanding pain mechanisms. The gate control theory also stated the influence of psychological or cognitive factors on pain experience. The gate control theory proposes that there is a gate, possibly located at the spinal cord level between the peripheral nerves and the brain, influencing the transmission of potentially painful signals to the level of conscious awareness. There is no pain or decreased intensity of pain when the gate is closed. The following situations make the gate closed: (a) the large diameter nerve fibers are activated by some factors such as skin stimulation; (b) inhibitory impulses from the brainstem which are caused by some factors such as distractions or guided imagery; and (c) inhibitory impulses

from the cerebral cortex and thalamus which are caused by some factors such as anxiety reduction (McCaffery & Beebe, 1994). There is pain when the gate is open. The following situations make the gate open: (a) the small diameter nerve fibers are activated by some factors such as tissue damage; (b) facilitatory impulses from the brainstem such as monotonous environment; and (c) facilitory impulses from the cerebral cortex and thalamus such as fear of pain (McCaffery & Beebe).

Postoperative Pain

Postoperative pain is a form of acute pain and is due to the surgical trauma with an inflammatory reaction and initiation of an afferent neuronal barrage. Peeters-Asdourian (2002) explained the mechanisms of postoperative pain as follows. The surgical incision destroys the local tissues involving nerve endings and activates specific nociceptors (pain receptors) and the free nerve endings. Then it leads to the release of inflammatory mediators such as bradykinin, serotonin, and histamine. These inflammatory mediators contribute to peripheral sensitization which is an amplification of noxious pain signals and manifested by hyperalgesia in clinic. Then these painful signals are transmitted to the dorsal horn of the spinal cord in an amplified fashion and are increased in duration. Finally, the periphery signals entering the central nervous system from the periphery will be increased in amplitude and duration. This phenomenon is called “wind up” or “central sensitization”.

Impacts of Unrelieved Postoperative Pain

Postoperative pain is the most common type of acute pain. Unrelieved postoperative pain leads to physiological and psychological problems and influences socioeconomic aspects.

Unrelieved postoperative pain leads to many negative physiological effects in some systems especially in elderly since the physical function is declined with the increased age and elderly have an increased vulnerability to stressors. In terms of the respiratory system, unrelieved pain decreases the effective cough reflex and increases atelectases, pneumonia, and hypoxemia (Breivik, 2002; Weetman & Allison, 2006). In addition, unrelieved pain can cause respiratory failure in elderly (Aubrun & Marmion, 2007; Craig, 1981). In the cardiovascular system, unrelieved pain increases heart rate and blood pressure, myocardial oxygen demand and ischaemic events (Breivik), and the risk of developing deep vein thrombosis (Aubrun & Marmion). Moreover, unrelieved pain can easily cause myocardial infarction in elderly (Aronow, 2003; Aubrun & Marmion). In the gastrointestinal system, unrelieved pain delays gastric emptying time leading to nausea and vomiting, decreases gut motility leading to ileus, and causes poor nutrition leading to delayed tissues healing (Aubrun & Marmion; Macintyre & Ready, 2002; Weetman & Allison). In the urinary system, unrelieved pain easily causes urinary retention (Breivik).

Psychologically, unrelieved pain increases anxiety, sleep disorders, and distress (Macintyre & Ready, 2002). In addition, unrelieved pain or medications used

to treat pain may cause confusion or delirium in elderly (Aubrun & Marmion, 2007). The unrelieved postoperative pain not only influences physiological and psychological aspects but also influences socioeconomic aspects such as delaying rehabilitation and increasing hospital cost (Brevivik, 2002).

Postoperative Pain Management

Postoperative pain management includes pharmacologic and nonpharmacologic management. These two approaches should be combined together in order to achieve effective pain management (Brunner & Suddarth, 2000; Mackintosh, 2007).

Pharmacologic Management

In the pharmacologic pain management, analgesic agents are the most commonly used agents to treat pain. Three general classes of analgesic agents are opioids, nonsteroidal anti-inflammatory drugs (NSAIDs), and local anesthetics (Brunner & Suddarth, 2000; Pasero & McCaffery, 1996). When managing acute postoperative pain, we need to start with strong opioid-based analgesics then down to non-opioid analgesics (Mackintosh, 2007).

The first class of analgesic agents is opioids. The mechanism of opioids to relieve pain is that opioids block the transmission of pain by acting at opioids receptor sites in the brain and spinal cord (Pasero & McCaffery, 1996). Morphine is

the most commonly used opioid for the severe postoperative pain relief (Mackintosh, 2007) and it has minimal side effects. The common side effects are constipation, sedation, nausea, and vomiting. Although the most serious side effect is respiratory depression, only appears in rare cases (Mackintosh). Moreover, morphine is a safe and an effective drug and its benefits outweigh its side effects (Mackintosh; Pasero & McCaffery). Morphine can be used in patient controlled analgesia (PCA). PCA pumps usually use morphine or morphine in combination with an anti-emetic drug to administer a small patient controlled intravenous dose. The majority patients using PCA can get a better pain relief (Chen, Chui, Ma, & Gin, 2001). However, clinicians need to teach the patients how to use it and also need continuous assessment, ensuring that the device works smoothly and pain relief is effective (Mackintosh).

The second class of analgesic agents is NSAIDs. NSAIDs decrease pain by inhibiting pain mediators such as prostaglandins, bradykinins, and histamines at the site of incision or injury (Pasero & McCaffery, 1996). NSAIDs are also effective in combination with opioids for treating moderate to severe postoperative pain (Pasero & McCaffery). However, some patients can not tolerate the side effects of NSAIDs. These side effects may involve upper gastrointestinal disturbances and renal impairment. In addition, NSAIDs must be used carefully in elderly (Pasero & McCaffery).

The third class of analgesic agents is local anesthetics. Local anesthetics work by blocking sensory inflow at the dorsal horn (Pasero & McCaffery,

1996). The local anesthetics can be used in two ways (Mackintosh, 2007). The first way is used in infiltration of wound during the final stages of surgery. This way can maintain an anesthetic effect on the wound area for some hours after surgery. The second way to use local anesthetics is by administrating anesthetics through an epidural catheter directly to the nerve root. By this way, the risk of sedation and respiratory depression can be reduced (Pasero & McCaffery) because it delivers analgesics directly to the site of their action in the central nerves system (CNS) and fewer drugs are likely to reach the supraspinal structures.

In clinic, balanced analgesia is a multimodal approach to manage postoperative pain. It can maximize pain relief and minimize potential adverse effects which usually come from high dose of any single agent (Pasero & McCaffery, 1996). A balanced analgesia for severe postoperative pain can include an opioid and a local anesthetic by given epidurally and NSAIDs by oral or parenteral (Pasero & McCaffery).

Nonpharmacologic Interventions

Although pain medication is the most effective method to relieve pain, nonpharmacologic interventions can also provide postoperative pain relief and reduce the dose requirement for analgesics (Rosenquist & Rosenberg, 2003). There are some nonpharmacologic interventions which have been examined for their ability to relieve pain such as cutaneous stimulation and massage, ice and heat therapies, distraction,

guided imagery, and relaxation techniques (Brunner & Suddarth, 2000). The mechanisms of these nonpharmacological interventions are different from each other. However, several nonpharmacological pain relief strategies, including rubbing the skin and using heat and cold, are based on the gate control theory. Many studies are focusing on using the relaxation techniques to relieve pain because they have minimal side effects and enable patients to learn self-care which is an important strategy for effective pain management. The relaxation techniques usually involve abdominal breathing at a slow and rhythmic rate or muscle relaxation. The relaxation techniques can produce the relaxation response by a quiet mind and a decrease in sympathetic stimulation of the hypothalamus (Roykulcharoen & Good, 2004) and relaxation also can increase endogenous opioid secretion which can reduce pain (Heffline, 1990).

Therefore, pharmacological management and nonpharmacological management should be used together to achieve maximum pain relief for patients. Besides the ways of managing pain mentioned above, the nurse should also need to adequately assess pain severity in heterogeneous population and monitor physiological and psychological changes associated with pain (Dunwoody, Krenzischek, Pasero, Rathmell, & Polomano, 2008).

Factors Related to Pain Experience

There are many factors influencing pain experience. According to the literature review, operation type, age, gender, educational level, cultural background,

previous experience, and patient's belief about pain can influence the pain experience. These factors may increase or decrease the person's perception of pain, evaluation of pain, and responses to pain.

Operation Type

Certain procedures have been recognized by clinicians as being very painful including thoracotomy, total knee replacement, and abdominal aortic aneurysm resection (Lynch et al., 1997). In a series of general surgical patients, Loan and Morrison (1967) found that the percentage of patients requiring pain relief depended on the site of operation. They found that more than 60% of patients undergoing thoracotomy and upper abdominal operations required pain medication compared with 12% undergoing neck operations. Chung and Lui (2003) studied the patient's level of postoperative pain in different operation approaches including thorax, abdomen, muscle-skeletal, ophthalmic, and dermatology. They found that the patients undergoing abdominal and musculo-skeletal surgeries had more severe pain intensity and less pain intensity was found in the subjects with ophthalmic operations.

Age

Pain perception does not change significantly with ageing (Chakour, Gibson, Bradbeer, & Helme, 1996). Pain assessment in older adults may be difficult because of the physiologic, psychosocial, and cognitive changes which accompany

ageing (Brunner & Suddarth, 2000). One study found that elderly patients self-administered less opioids than younger patients in the first two postoperative days (Gagliese & Katz, 2003). In addition, the elderly might be more reluctant to report painful stimuli than younger people because the elderly adopted a more conservative response bias when experiencing noxious stimuli (Gagliese & Melzack, 2006). However, elderly report pain more apparently when the stimulus is very intense and/or persists for long periods (Aubrun & Marmion, 2007).

Gender

It is generally accepted that gender influences pain experience (Berkley & Holdcroft, 1999). One study by Keogh and Herdenfeldt (2002) showed that pain coping instructions had different effects on the pain experience of males and females. Compared to females, males showed less negative pain responses when focusing on the sensory component of pain (e.g., increased threshold, tolerance, and lower sensory pain). However, when compared to sensory focusing, emotional focusing was found to increase the affective pain experience in females that females reported more pain experience, more severe levels of pain, more frequent pain, and longer duration of pain (Unruh, 1996). Rollman, Lautenbacher, and Johnes (2000) revealed that females had lower pain thresholds and pain tolerance to a wide range of noxious stimuli. Chung and Lui (2003) studied the level of postoperative pain on patients. They found that female reported higher levels of current pain intensity and worst pain intensity compared with

males. However, it is yet to be determined whether the higher reported pain intensity is due to the higher sensitivity in somatic responses to pain stimuli after operation, or due to that it is more socially acceptable for females to express pain.

Educational Level

The persons with higher level education are expected to be more aware with the things related to their health and lives. Chung and Lui (2003) studied the patient's level of postoperative pain and found that educational level influenced the patient's perception of pain that older subjects and those with a lower educational level perceived less pain. This may be due to that the older and less well-educated persons were not adequately informed about their rights to receive pain relief. Ward et al. (1993) found that the misconceptions and barriers to effective pain management were found to be more in the subjects with fewer years' education.

Cultural Background

Cultural background which is an important aspect of sociocultural dimension of pain influences how the person perceives (Douglas, 1999; Lasch, 2000) and responds to pain (Lasch; Martinelli, 1987). Cultural background has been also shown as an important factor that influences pain behavior and expression (Martinelli; McGuire, 1992). For example, a person from one culture is taught what stimulus is expected to be painful and what responses to pain are acceptable. Therefore, people

from different culture who have the same pain intensity may respond to pain differently. One study investigated how Taoism/energy, Buddhism, and Confucianism influenced Chinese patients' perspectives on cancer pain (Chen, Miaskowski, Dodd, & Pantilat, 2008). They found that within the beliefs of Taoism/energy, the blockage of Qi/blood must be removed in order to relieve pain; a Confucian believes that he/she should endure the pain and not report it to a clinician until the pain becomes unbearable; within the beliefs of Buddhism, the pain can end only by following the eight right ways such as right view and right action.

Previous Experience

The person is more frightened about subsequent painful events if he/she has more pain experience. This may be due to that once a person experiences severe pain, he/she knows just how severe it can be (Brunner & Suddarth, 2000). Therefore, the person may have no fear of severe pain if he/she never experiences severe pain before.

Patient's Belief About Pain

Some studies found that the patient's belief about pain influenced the patient's pain report. One qualitative study explored pain experience and beliefs of Chinese patients undergoing surgery (Wong & Chan, 2008). They found that there were some beliefs in the patients which might be the barriers to pain report. These

beliefs included: (a) lack of control over pain. They felt that nothing could be done to control their pain even with analgesics; (b) pain is a negative signal. The pain intensity they experienced was a negative sign to their present problem and future health; (c) worry about “Shan”. “Shan” is a term used in Chinese traditional medicine which means that the drug would cause dizziness, nausea, and vomiting. Thus, the patient tried to bear their pain in order to avoid the side effects of drugs; (d) being a good patient. They worried about that they might be perceived by nurse as too demanding if they asked for pain relief frequently; and (f) passive coping which involves avoiding thinking about the pain, avoiding negative thoughts, stoically tolerating pain, and avoiding any movement of the affected limb. Another study explored the concerns related to reporting pain and using analgesics among postoperative patients (Tzeng, Chou, & Lin, 2006). They found that the following beliefs made the patients hesitated to pain report and use analgesics: (a) time interval, the belief that pain medication can be given only at specified time intervals and one can not ask for medication until that interval has passed; (b) fear of tolerance; (c) fear of wound healing inhibition; (d) fear of distracting the physician from treating the disease; (e) fatalism, the belief that pain is an inevitable consequence of disease; and (f) a desire to be a good patient.

Measurement and Pain Assessment Measures

Definition of Measurement and Measurement Theory

Measurement is crucial to science, however, it is also a very abstract concept to define and understand in social sciences. Waltz, Strickland, and Lenz (2005) stated the definition of measurement as “the process of using a rule to assign numbers to objects or events which represent the kind and/or amount of a specified attribute possessed” (p.43). Measurements focus on the attribute which is variable and not constant and measurements take on different values for different objects that termed variables in scientific language (Waltz et al., 2005). The results of measurement, obtained by using the precise procedure, are commonly expressed by numbers which are less vague than words and can communicate information more accurately (Polit & Beck, 2004; Waltz et al.). Phenomenon to be measured is so abstract that cannot be identified totally as either objects or events. Thus, Carmines and Zeller (1979) stated that measurement is the process to link abstract concepts to empirical indicators and concerns the crucial relationship between the empirically grounded indicator which is the observable response and underlying unobservable concept.

Classical measurement theory is the foundation of a model for assessing random measurement. Random error, also termed chance error, is caused by chance factors that confound the measurement of any phenomenon (Waltz et al., 2005). The basic principle of the classical measurement theory comes from the assumption that

random error is an element that must be considered in all measurement (Waltz et al.). Social scientists try to eliminate as much random error from their measurements as possible, but the measurements still contain a limited amount of random error. This theory has a basic formulation that every observed score is made up of a true score and an error score. Waltz et al. explained this formulation as follows. For the true score, it is the true or precise amount of the attribute possessed by the object or event being measured. One implication we can get from the classical measurement theory is that systematic errors can become part of true score and affect validity but not reliability. For the error score, it reflects the influence that random error has on the observed score. Therefore, random error directly influences the reliability of the measurement. The influence of the random error on the observed measurement is called the error of measurement. When the phenomenon is measured in an infinite number of times, the distributions of random error can be expected to be normally distributed. The standard deviation of error scores is termed the standard error of measurement. The more the obtained scores vary about the true score, the more measurement error there is. Therefore, the measurement procedure is more reliable when the standard error of measurement is smaller (Waltz et al.).

Psychometric Properties of Instrument

Instrumentation is the process of selecting or developing tools and methods appropriate for measuring an attribute or characteristic of interest (Waltz et al.,

2005). The tool is called instrument. In order to quantify an instrument, the psychometric properties of the instrument need to be assessed. Validity and reliability are two major criteria for assessing the psychometric properties of an instrument (Polit & Beck, 2004).

Validity of Instrument

Definition of validity. Carmines and Zeller (1979) stated “an indicator of some abstract concept is valid to the extent that it measures what it purports to measure” (p.12). Validity focuses on the crucial relationship between concept and indicator and one validates not the measuring instrument itself but the use to which the instrument is put (Carmines & Zeller; Nunnally & Bernstein, 1994).

Validity and nonrandom error. Nonrandom error, also termed systematic error, has a systematic biasing influence on measurement procedures (Waltz et al., 2005). Nunnally and Bernstein (1994) described the features of systematic error as followings: (a) systematic error can affect all observations equally and be a constant error; (b) systematic error can affect certain types of observations differently than others and be a bias. For example, a miscalibrated thermometer that always reads two degrees too low illustrates a constant error in the physical sciences. Take for another example, if the FPS for measuring pain is more sensitive to some irrelevant attributes

such as anger or sourness, the error would be a bias. The systematic bias in all results obtained by the tool or method influences the extent to which the attribute of interest is actually measured (Waltz et al.). Thus, the occurrence of nonrandom error is the central threat to validity.

Types of validity. There are four types of validity including content validity, face validity, criterion-related validity, and construct validity.

1. Content validity

For content validity, Carmines and Zeller (1979) stated “content validity depends on the extent to which an empirical measurement reflects a specific domain of content” (p. 20). Polit and Beck (2004) stated another definition: “content validity concerns the degree to which an instrument has an appropriate sample of items for the construct being measured” (p. 423). Content validity is important for affective measures and cognitive measures (Polit & Beck). Content validity for cognitive measures, such as knowledge test, refers to the extent to which the questions on the test represent the universe of the questions on the topic. In the development of affective measures, content validity is also important and the new instrument needs to start with a thorough conceptualization of the construct in order to cover the entire content domains.

Although there are no completely objective methods to evaluate content

validity of an instrument, the panel consisting of at least three experts to evaluate and document the content validity of new instruments is increasingly common (Polit & Beck, 2004). Waltz et al. (2005) stated the process of assessing content validity in different situations as follows. In the situation that only two experts are employed, the content validity index (CVI) is used to assess the degree of agreement between the experts. To compute the CVI, two content experts are given the objectives and items. Then they are asked to independently rate the relevance of each item to the objective(s) by using a 4-point rating scale. In this rating scale, 1 means not relevant, 2 means somewhat relevant, 3 means quite relevant, and 4 means very relevant. The CVI equals the proportion of items given a rating of quite relevant and very relevant by both experts. A CVI score of .80 or better indicates good content validity. In the situation that more than two experts are employed, the alpha coefficient rather than CVI is more likely to be used as the index of content validity. A coefficient of 0 means lack of agreement between the experts. A coefficient of 1.00 indicates complete agreement. This complete agreement does not mean that the same rating is assigned by all experts, but means that the relative ordering from one expert matches the relative ordering from the other experts.

2. Face validity

Polit and Beck (2004) stated “face validity refers to whether the instrument looks as though it is measuring the appropriate construct” (p. 423). The

instrument has face validity if a cursory inspection of the instrument appears to measure what the test constructor claims it measures (Waltz et al., 2005). Although face validity should not be considered to provide evidence for validity, response rates of the participants may increase in an evaluation if the instruments being used have face validity (Polit & Beck; Waltz et al.). Thus, face validity by itself never provides sufficient evidence to establish validity, however, we still need to examine it making us to be confident in the decisions we make based on the tests scores (Nunnally & Bernstein, 1994).

According to Mc Donald (2002), face validity refers to what a test appears to measure, the appearance of the test coincides with its use, and face validity also helps to keep the motivation of the test takers high. Therefore, in this study, face validity of the five pain intensity scales would be assessed by three aspects including preference, simplicity, and accuracy and these three aspects would be assessed separately.

Firstly, preference was assessed by using the Scale Preference Questionnaire (SPQ). On the SPQ, five pain intensity scales were presented and the participants were asked to rank order from 1 to 5 when 1 = most preferred and 5 = least preferred. Secondly, simplicity was assessed by using the Scale Simplicity Questionnaire (SSQ). On the SSQ, five pain intensity scales were presented and the participants were asked to rank order from 1 to 5 when 1 = simplest and 5 = least simple. Lastly, accuracy was evaluated by the researcher using the Scale Accuracy

Checklist (SAC). The SAC was developed by the researcher from a literature review to assess the accuracy of the pain scales. In the SAC, there were six types of error: (a) ratings outside the scale range (ratings between scale units or above/below the end-points of the scale); (b) no rating on the scale; (c) more than one rating on a single scale; (d) a range of ratings on one scale; (e) response falling between two numbers, words, or facial expressions; and (f) mistake in ordinal understanding of the scale (Peters et al., 2007). The meaning of mistake in understanding of the scale is explained as follows. The participants need to rate four assessments for each scale including current pain and daily retrospective worst, least, and average pain in the past 24 hours in this study. If the participants understand the scale and its labels rightly, the worst pain score should be rated as higher than or at least equal to the average pain score and the average pain score should be rated as higher than or equal to the least pain score. Therefore, if the worst pain score minus average pain score or average pain score minus least pain score yields a negative result, it would be counted as a mistake in ordinal understanding of the scale. For the above errors, only one type of error was counted per participant for each scale. For example, if one participant does not have ratings in all four pain intensity questions on the NRS, this is counted as one error for the NRS. Thus, on the SAC, there were six types of errors and the accuracy of the scale was evaluated by the number of subjects with accurate response (without any errors) and the number of the subjects with inaccurate response (with any errors). In conclusion, for assessing face validity, the higher number of subjects indicating the

most preferred scale, the simplest scale, and the higher number of subjects with accurate response of the scale indicated the higher face validity.

3. Criterion-related validity

Nunnally and Bernstein (1994) stated that criterion-related validity concerns “using an instrument to estimate some criterion behavior that is external to the measuring instrument itself” (p. 94). Generally, the aim of testing the criterion-related validity is that the investigator wants to know the extent to which performance on an important criterion can be estimated by using information from a less costly and more easily obtained measure (Waltz et al., 2005). There are two types of criterion-related validity including concurrent validity and predictive validity. Concurrent validity is assessed by correlating a measure and the criterion within a short period of time when the criterion exists in the present; however, predictive validity focuses on a future criterion which is correlated with the relevant measure (Carmines & Zeller, 1979). The difference between these two types of validity is that whether they concern the current or future existence of the criterion variable rather than the logic and procedures of validation (Carmines & Zeller). When testing the criterion-related validity, three issues need to be considered (Waltz et al.). The first issue is that the criterion should be a high-status operationalization of the same construct that the predictor is trying to assess and not an operationalization of some other construct. The second issue is that the criterion and/or the predictor must

demonstrate sufficient reliability or else may lead to a reduced criterion-related validity coefficient. The third issue is related to an assumption underlying criterion-related validity procedures that the nature of the target population is relatively static. Therefore, when assessing concurrent validity, the measure being tested for validity and the related criterion need to be given to the subjects within a short period of time. Then the results on these two measures are compared by using the appropriate correlation or regression procedure. Finally, the results of this comparison indicate the predictor measure's ability to predict present standing on the criterion.

In this study, since postoperative pain would interfere with patients' functioning, the modified Pain Interference Scale (PIS) which was under a Chinese version of the Brief Pain Inventory (BPI-C) was used as the criterion to assess the concurrent validity of each pain scale. The modified PIS was administered to the patients within a short period after the patients finish rating current pain and daily retrospective worst, least, and average pain by using the five pain scales. As the rating of recalled worst pain might better reflect the overall experience of pain and its impact on function (Shi, Wang, Mendoza, Pandya, & Cleeland, 2009), this study would choose the scores on worst pain rating to assess the construct validity. Therefore, concurrent validity was assessed by calculating Pearson product-moment correlation coefficient between the scores of the modified PIS and scores of each pain intensity scale by using the recalled worst pain during the past 24 hours. The higher is the correlation, the higher the concurrent validity.

4. Construct validity

Carmines and Zeller (1979) stated “construct validity is concerned with the extent to which a particular measure relates to other measures consistent with theoretically derived hypotheses concerning the concepts (or constructs) that are being measured” (p. 23). Just as Waltz et al. (2005) stated, “the primary concern in assessing construct validity is the extent to which relationships among items included in the measure are consistent with the theory and concepts as operationally defined” (p. 156). Construct validity can be determined by several ways but it always involves logical analysis and tests predicted by theoretical considerations (Polit & Beck, 2004). Those several ways are known-group technique, multitrate-multimethod matrix (MTMM), and factor analysis (Polit & Beck).

4.1 Known-group technique

Known-group technique is an approach that the instrument is administered to groups expected to differ on the critical attribute because of some known characteristic (Polit & Beck, 2004). For example, when validating an instrument of fear of postoperative pain, we can choose the surgical ward patients who are going to have operations. Then we contrast the scores of the patients without surgery history and the patients who had experienced surgery before. We would expect that the patients who had experienced surgery before would be more fear about postoperative pain than the patients without surgery history. Therefore, we might

question the instrument's validity if such differences do not emerge. We would not necessarily expect large differences as some patients without surgery history would have more fear of pain, and some patients who had experienced surgery before would express less fear. On the whole, however, we would anticipate differences in the average fear scores between these two groups.

4.2 The multitrate-multimethod matrix (MTMM)

The multitrate-multimethod matrix (MTMM) is a significant construct validation procedure which involves convergent validity and discriminant validity. Convergent validity refers to the correlations between different methods that measure the same trait (Polit & Beck, 2004). If the correlations are high enough, they will provide evidence for convergent validity. Validation is generally convergent because it is concerned with demonstrating that two different methods measuring an attribute lead to similar results (Nunnally & Bernstein, 1994). Discriminant validity focuses on the instrument's ability to differentiate the construct from other similar constructs (Polit & Beck).

There is an example for helping to explain the MTMM. Suppose a nurse had two scales designed to measure anxiety level including one self-report scale and one observation scale. She also had another two scales designed to measure pain intensity level including one self-report scale and one observation scale. Each of these four scales is administered to every participant at the same time. The reliability of each

instrument is then determined by using an index of internal consistency (alpha/KR-20/KR-21) and the correlations between each pair of forms are computed. If the results of reliability are sufficiently high, the procedure continues; if not, the procedure terminates as reliability is a prerequisite for validity. The most direct evidence of convergent validity comes from the correlation between two different methods measuring the same trait. Therefore, if the self-report scale of anxiety and the observation scale of anxiety correlate high with each other, these two scales are proved to have convergent validity. The evidence of discriminant validity comes from the following two requirements. For one requirement, the correlation between two different methods measuring the same trait should be higher, in absolute magnitude (refers to the value without a plus or minus sign), than the correlations between measures that have neither method nor trait in common. For another requirement, the coefficient between two different methods measuring the same trait should be greater than the coefficients between measures of different traits by a single method. Generally, if these two requirements are obtained, it will provide evidence for discriminant validity. The situation that can fit with the MTMM approach is usually not available. Interpreting the pattern of coefficients is also complex in MTMM and the evidence is seldom clear. However, MTMM is still a valuable tool for assessing construct validity even the full model is not feasible (Polit & Beck, 2004). Therefore, it is acceptable that many researchers focus only on convergent validity.

4.3 Factor analysis

Factor analysis is one way to examine the correlations among the factors or items which may have underlying relationships between each other (Polit & Beck, 2004). For example, the investigator has designed an instrument to assess various dimensions of a phenomenon and this instrument is based on a conceptual framework. If he/she wants to empirically justify these dimensions, factor analysis will be a useful method to assess construct validity of this instrument.

Based on the feasibility of the above methods that used to assess construct validity, the researcher would use convergent validity to assess the construct validity of the pain scales in this study. The recalled worst pain score of the VDS was used as the gold standard score to be correlated with the recalled worst pain scores on the NRS, the FPS, the CAS, and the BS-21 because the VDS is thought to be easier for older adults to understand and prior psychometric validation (Closs et al., 2004; Gagliese et al., 2005; Peters et al., 2007). Thus, convergent validity was assessed by calculating Pearson product-moment correlation coefficient between the VDS and each of the four pain scales. Since the NRS is another tool that is easily to understand and is recommended for pain assessment in elderly patients including those with none to mild or moderate CI from a study of literature review (Hadjistavropoulos et al., 2007), therefore, if the correlation between the NRS and the VDS is high enough, it would provide evidence for convergent validity of both the NRS and the VDS. For other three scales, the higher is the correlation, the higher the convergent validity of

that scale.

Reliability of the Instrument

Definition of reliability. For reliability, Waltz et al. (2005) stated “a measurement tool is reliable for a particular subject population to the extent to which it yields consistent results on repeated measurements of the same attribute” (p.59). Reliability concerns a particular property of empirical indicators that the degree to which they provide consistent results across repeated trials (Carmines & Zeller, 1979). Reliability is directly related to the variance of obtained scores (Nunnally & Bernstein, 1994). Therefore, the measuring instrument is more reliable if the results given by repeated measurement of the same phenomenon are more consistent.

Reliability and random error. Random error, totally unsystematic in character, is used to describe those chance factors that confound the measurement of any phenomenon (Carmines & Zeller, 1979). The random error does not directly influence the meaning of the measurement but does directly influence the precision with which the characteristic of interest is being measured (Carmines & Zeller). The occurrence of random error is the central threat to reliability which concerns the consistency of results across repeated measurements (Carmines & Zeller; Waltz et al., 2005). Nunnally and Bernstein (1994) stated that reliability is free from random error

only when the researcher considers how repeatable observations are: (a) when different persons take measure; (b) with alternative instruments intended to measure the same thing, and (c) when incidental variation exists in the conditions of measurement. This last point implies that the contents on multi-item test should be homogenous and internal consistency should be high among components of the overall measure (Nunnally & Bernstein). In addition, the researcher needs to consider that measurement must be stable over time.

Types of reliability. There are four basic methods for estimating the reliability of empirical measurement. These are the test-retest reliability, the alternative-form method, interrater reliability, and internal consistency reliability.

1. Test-retest reliability

Test-retest reliability is one way to estimate the reliability of empirical measurements in which the same test is given to the same people after a period of time. The reliability is equal to the same test obtained at two points in time. When data are measured at interval level, the Pearson product-moment correlation coefficient is used to estimate reliability. When data are measured at nominal or ordinal level, a nonparametric statistics, such as Chi-square or Spearman rank correlation coefficient, is used. Test-retest reliability focuses on the instrument's susceptibility to extraneous factors over time. There are mainly three issues that need to be considered when using

test-retest method. Firstly, the reliability will be less than perfect because of the instability of measures taken at multiple points in time (Carmines & Zeller, 1979). Secondly, the reliability is influenced by the memory interference. Memory interference refers to the fact that the second administration will be influenced by the subjects' memory of initial responses (Carmines & Zeller). In order to avoid the memory interference, the second administration should occur approximately two weeks after the first administration (Waltz et al., 2005). Lastly, Nunnally and Bernstein (1994) stated another issue of test-retest reliability is that a measure that has no internal consistency may be quite stable over time.

In this study, test-retest method was used to assess the reliability of the pain intensity scales and the issues about using test-retest method were considered. Firstly, the reliability would be less than perfect because of the instability of measures taken at multiple points. Thus, in this study, the subjects were asked to recall the vividly remembered painful experience ever felt in their lives rather than the unstable postoperative pain. Secondly, the reliability is influenced by the memory interference. Therefore, a 2-week period is advisable to complete both tests. However, the researcher also needed to consider the memory deficits as the elderly with mild CI were included in this study. Chibnall and Tait (2001) compared pain assessment scales in elderly without and with CI. They found that test-retest reliability for the elderly with CI group dropped greatly when the reliability was evaluated over a 2-week period. Even across a single week, the reliability estimates were considerably lower in the

elderly with CI group than in the elderly without CI group. However, when the period was restricted to 3 days in the elderly with CI group, the reliabilities approximated the reliabilities obtained in the elderly without CI group for 7-day period. Thus, 3-day period for test-retest reliability is reasonable in the elderly with CI group. Another study (Taylor & Herr, 2003) also supports this point. They compared pain scales in African American elderly without and with CI and found that pain scales demonstrated acceptable test-retest reliability in the elderly without CI group and to a lesser degree in the elderly with CI group at a 2-week interval ranged from .52 to .83 in both groups. Furthermore, Taylor, Harris, Epps, and Herr (2005) compared pain scales in elderly without and with CI. They found that test-retest reliability over a 2-week interval was acceptable (.67 - .85) in elderly without CI group but unacceptable in elderly with CI group (.26 - .67) and proposed that the stability issue must be considered in the elderly with CI group. Therefore, based on the above studies, the test-retest reliability in this study would be assessed over a 3-day period.

In short, test-retest reliability of the pain scales in this study would be examined by the way that subject was asked to rate the intensity of any vividly remembered painful experience ever felt in his/her life by using the five pain scales on the preoperative day and 3rd postoperative day. Test-retest reliability was assessed by correlating the scores that obtained twice. The higher is the correlation, the higher the test-retest reliability.

2. Alternative-form method

Alternative-form method is that the alternative form of the same test rather than the same test is administered on the second testing (Carmine & Zeller, 1979). These two forms of the test are intended to measure the same thing. The obvious advantage of this method compared with test-retest method is that it reduces the memory interference. However, the limitation of alternative-form method is the practical difficulty of constructing alternative forms that are parallel (Carmine & Zeller).

3. Interater reliability

Interater reliability refers to the consistency of performance among different raters in assigning scores to the same responses or objects (Waltz et al., 2005). Interater reliability is very important when observational measures are used and subjective measures are included, such as free responses requiring categorizing, case studies, and essays (Waltz et al.).

4. Internal consistency reliability

An instrument may be said to be internally consistent to the extent that its items measure the same trait (Polit & Beck, 2004). Therefore, the internal consistency reliability can be assessed only for the instrument with multiple items. The split-half method, alpha coefficient, Kuder-Richardson formula 20 (KR-20), and

Kuder-Richardson formula 21 (KR-21) can be calculated to estimate internal consistency reliability. These methods based on internal consistency are difficult to use with speed tests (Nunnally & Bernstein, 1994). For all new measurement methods, internal consistency should be assessed even if other estimates of reliability are also necessary (Nunnally & Bernstein).

4.1 Cronbach's alpha

Alpha represents the extent to which performance on any one item on an instrument is a good indicator of performance on any other item in the same instrument (Waltz et al., 2005). A higher alpha value usually indicates that all the items in the test measure just one attribute. The researcher needs to know that when testing an instrument that measures more than one attribute (e.g., those with subscales), alpha should be tested for each subscale rather than the whole instrument. The researcher also needs to be careful when interpreting the alpha value. According to Waltz et al. the following issues may affect the alpha value. The issues that come from the formula for determining the alpha coefficient are: (1) the more items are included, the higher alpha value is obtained; and (2) the alpha value also depends on the total test variances and the shape of the resulting distribution of test scores. Another issue is that when most respondents can not complete the test, a higher alpha may be obtained. Therefore, when less than 85% of the subjects respond to all items on the test, the alpha value should not be used to estimate internal consistency reliability.

4.2 Kuder-Richardson formular 20 (KR-20) and Kuder-Richardson formular 21 (KR-21)

KR-20 or KR-21, simply a special case of alpha, is to estimate the reliability of scales composed of dichotomously-scored items (Carmines & Zeller, 1979; Waltz et al., 2005). Dichotomously scored test is that when each item in the test is scored one if correct and zero if incorrect. KR-20 is used when the item difficulty levels can not be assumed to be the same. When one can assume that the difficulty level of all items is same, then KR-21 may be employed (Waltz et al.).

4.3 Split-half method

For split-half method, items on a scale are split into two halves and scored independently. The correlation coefficient for scores on the two-half tests gives an estimate of the test's internal consistency (Polit & Beck, 2004). If the items on the two-half tests measure the same attribute, the reliability coefficient should be high. Polit and Beck proposed the following three issues when using split-half method. Firstly, the correlation coefficient computed on split-half tends to underestimate the reliability of the entire scale. Second, longer scales are more reliable than shorter ones. Third, different reliability estimates can be obtained with different splits. In addition, split-half method is likely to lead to misleading estimates when items on the test are ordered in terms of difficulty (Nunnally & Bernstein, 1994).

Relationship of Reliability and Validity

It is important to understand the relationship of reliability and validity. Nunnally and Bernstein (1994) stated “reliability is necessary but not sufficient to validity” (p.214). They explained that consistency of results does not necessarily mean that the measurement accurately measures what it is purports to measure. However if a measurement is unreliable, the results are not meaningful which means that the indicator cannot represent the attribute intended to measure. For example, the number on the ruler shows that the length of the ruler is 10 cm but in fact the length of the ruler is only 9 cm. Thus, when this ruler is used to determine the length of a table, it overestimates the length of the table by 1 cm for every 10 cm. The results will be quite consistent on repeated measurements although they are obviously incorrect. In short, this ruler will provide a quite reliable but totally invalid indication of the length.

Interpretation of Validity and Reliability Coefficients

In order to interpret the validity and reliability coefficients in this study, the criterion for interpreting the correlation coefficient and the issues when interpreting reliability and validity coefficients are reviewed.

The correlation coefficient is used as a tool for quantitatively describing the magnitude and direction of a relationship between two variables (Polit & Beck, 2004). Correlation coefficients summarize how perfect the relationships are. The correlation coefficient ranges from -1.00 through .00 to +1.00. Coefficients running

from .00 to -1.00 indicate negative relationships. When two variables are totally unrelated, the correlation coefficient equals zero. Coefficients ranging from .00 to +1.00 indicate positive relationships. According to Munro (2001), interpretation the correlation coefficient depends on the situations. Alternate forms of a test should be measuring the same thing. Therefore, their correlation should be high. When the tests' results are used to make important decision, the correlations between two forms of the same test should be very high, approximately .95. However, when testing the relationships among different aspects of human behavior, a correlation of .50 is acceptable. Munro categorizes the strength of correlation as follows. .00 - .25 means little relationship if any, .26 - .49 means low, .50 - .69 means moderate, .70 - .89 means high, and .90 - 1.00 means very high. According to Polit and Beck (2004), for most psychosocial variables (e.g., stress and severity of illness), a correlation of .70 is high and correlations between such variables are typically ranging from .10 to .40.

The interpretation of reliability coefficients depends on what is being measured, the stage of development of the instrument, and the procedure used to estimate reliability (Jacobson, 2004). The reliability of physiologic measures is generally higher than that of attitudinal measures (Jacobson). Nunnally and Bernstein (1994) stated that an alpha coefficient of .70 is acceptable but modest for an instrument in the early stage of development and .80 is desirable for a more developed instrument and when the purpose is to compare groups. If the study is to make important decisions about individuals, a reliability of .90 is desirable (Jacobson). Reliability is increased by

longer test length, by speeded conditions in which all participants do not finish the test, by heterogeneous samples. However, the researcher needs to consider that very high reliabilities may indicate redundant items (Jacobson). Reliability coefficients should be calculated each time when an instrument is used, particularly if used on a different population. If an instrument consists of subscales, the reliability of each subscale must be assessed. The reliability of subscales is often lower than that of the total instrument because the subscales are shorter (Jacobson).

The validity coefficients are usually lower than those of reliability. Compared to reliability coefficients of .70 to .90, validity coefficients of .30 to .60 may be entirely satisfactory (Jacobson, 2004). Evidence for the validity of a test should be seen as a cumulative pattern (Jacobson). More positive results can give greater confidence for the validity. The validity evidence is specific to a use of the test scores for a certain group and purpose rather than to the instrument itself. Therefore, users should provide additional evidence of validity from their studies (Jacobson).

Based on the above information, for the interpretation of validity and reliability coefficients in this study, the researcher would use the following criteria.

Correlation Coefficients (r)	Level of relationship	Level of validity/reliability
> .70	High	Very good
.40 - .69	Moderate	Good
.20 - .39	Low	Fair
< .20	Very low	Poor

Pain Assessment Tools

Types of Pain Assessment Tools

Although there are many ways to category the pain assessment tools, two ways are commonly used. Firstly, the pain assessment tools can be categorized based on the characteristic of self-report, behavior, and physiology. Since pain is a subjective experience that can be only perceived directly by sufferers, self-report is the most reliable way for accurate assessment of pain (Pasero, 2003). However, behavioral assessment tools and physiological assessment tools are more suitable when self-report is not feasible in some situations. These situations may involve: (a) when the patient is under the influence of anesthetics or unconsciousness; (b) when lacking adequate verbal skills such as in neonates and children younger than three years; and (c) when assessing the mentally challenged patients (Ho, Spence, & Murphy, 1996). Secondly, the pain assessment tools can be categorized according to the dimensions of pain. Pain

is a multidimensional experience including physiological, sensory, behavioral, cognitive, affective and sociocultural dimensions (McGuire, 1992). By this way, the instruments measuring pain can be divided into two types: unidimensional assessment tools which measure only one dimension of pain and multidimensional assessment tools which measure two or more than two dimensions of pain. This review presents the pain assessment tools as the following three types: pain assessment tools based on self-report including the unidimensional and multidimensional tools, behavioral pain assessment tools, and physiological pain assessment tools.

Pain assessment tools based on self-report. Pain assessment tools based on self-report can be divided into two categories: unidimensional pain assessment tools and multidimensional pain assessment tools.

1. Unidimensional assessment tools based on self-report

For the unidimensional assessment tools, sensory dimension is most commonly used to measure pain (McGuire et al., 2004). Pain intensity which is under the sensory dimension is probably the most frequently assessed component of pain (Dahl, 1996; Gagliese & Melzack, 2006) because pain intensity is used to determine the effectiveness of pain management in clinic. There are many kinds of pain intensity scales and this study would focus on the Numeric Rating Scale (NRS), the Verbal Descriptor Scale (VDS), the Colored Analogue Scale (CAS), the Numerical Box-21

Scale (BS-21), and the Faces Pain Scale (FPS). The introductions and qualities of these five pain intensity scales are presented as follows.

The NRS is a line marked with 11 numbers (0 through 10) at equal intervals. 0 equals “*no pain*” and 10 equals “*worst pain*”. Participants were asked to select the number that represented their pain intensity. Many studies support the validity and reliability of the NRS. Gagliese et al. (2005) compared the VAS-horizontal (VAS-H), the VAS-vertical (VAS-V), the NRS, and the VDS in younger and older surgical patients. They found that the NRS was the preferred scale and had low error rates, and had higher face, convergent, divergent, criterion validity than other scales. In addition, its properties were not age-related. Moreover, according to a study of literature review, the NRS was easily to understand and recommended for pain assessment in elderly including those with none to mild or moderate CI (Hadjistavropoulos et al., 2007). Ware, Epps, Herr, and Packard (2006) supported the convergent validity of the NRS and its test-retest reliability at a 2-week interval in older minority adults including elderly with CI. Convergent validity of the NRS was also supported in both elderly without and with CI groups (Taylor & Herr, 2003). In the study of Closs et al. (2004), convergent validity of the pain scales was assessed by the correlation between the pain scales and convergent validity of the NRS was acceptable in the elderly patients including those with none to moderate CI.

There are many versions of the VDS. This study would use the VDS (0-4) as its high feasibility for pain assessment in postoperative elderly (Pesonen et al.,

2008). The VDS (0-4) consists of five numerically ranked descriptors. 0 equals “no pain”, 1 equals “slight pain”, 2 equals “moderate pain”, 3 equals “severe pain”, and 4 equals “unbearable pain”. Participants selected the word that represented their pain intensity. Many studies support the qualities of the VDS. The VDS had low error rates, good face and convergent and criterion validity in younger and older surgical patients (Gagliese et al., 2005). The VDS may be considered for pain assessment when the majority of subjects are elderly (Peters et al., 2007). The VDS was the most feasible pain scale for pain assessment in postoperative elderly patients (Pesonen et al., 2008). Convergent validity and test-retest reliability over a 2-week interval of the VDS were supported in older minority adults including elderly with CI (Ware et al., 2006). The VDS was the most sensitive and reliable tool in using with younger and older adults (Herr et al., 2004). The VDS was the most easily understandable and preferred scale and was suitable for all but patients with most severe CI; convergent validity of the VDS was high except in patients with severe CI (Closs et al., 2004). Moreover, according to a study of literature review, the VDS has been shown to be the most preferred and understandable tool and is recommended for pain assessment in illiterate patients and the patients with none to mild or moderate CI (Hadjistavropoulos et al., 2007). However, construct validity of the VDS was a little less strong in younger and older adults (Peters et al.).

The CAS is designed to assess pain intensity among children (McGrath et al., 1996) and it consists of two parts. One part that can be seen by the subjects is a

14.5 cm long triangular shape varying in width and hue. The width and hue range from 1 cm wide and light pink hue at the bottom labeled “*no pain*” to 3 cm wide and deep red hue at the top labeled “*most pain*”. On this part, there is also a marker and the subjects can select the scale position by sliding this marker from the bottom to the top. The other part is the back of the CAS that can not be seen by the subjects. This part shows the numerical value of the rating. Therefore, the subject’s score is the numerical value on the back of the scale which matches the selected scale position on the triangular shape. In this study, this scale would be modified since the CAS was presented on the questionnaire. Thus, two parts of the original CAS were integrated into one part in order to make it practical in this study. However, the modified CAS was still presented on a vertical format since Chinese people assess pain more accurately with a vertical version of the VAS than with the more commonly used horizontal version (Aun et al., 1986). Therefore, the modified CAS in this study consisted of a vertical triangular shape. This triangular shape, varying in width and hue, ranged from 1 cm wide and light yellow hue at the bottom to 2.5 cm wide and deep red hue at the top. These colors were selected based on the patients’ preference from a pilot study. The length of the triangular shape was 10 cm with anchors “0” and “*no pain*” at the bottom and anchors “10” and “*worst pain*” at the top. The participants were asked to mark a horizontal line on the triangular shape where best reflected their pain intensity and pain intensity was scored by measuring the vertical distance from the bottom to the patient’s mark. For the quality of the CAS, Scherder and Bouma (2000) compared the CAS, the FPS, and

the Facial Affective Scale (FAS) among an early and midstage of Alzheimer's disease (AD) population and nondemented elderly population. They found that CAS could be comprehended very well by nondemented elderly, early AD elderly, and midstage AD elderly, 100%, 100%, and 80%, respectively.

The BS-21 has a horizontal row of 21 boxes with numbers labeled from 0 to 100 in increments of five (Jensen et al., 1998). There are anchors "*no pain*" on the left extreme and "*worst pain*" on the right extreme. The participants selected the box that represented their pain intensity. Many studies support the qualities of the BS-21. Peters et al. (2007) compared the VAS, the VDS, the BS-11, and the BS-21 in younger and older patients. They found that the BS-21 had lowest incorrect responses and good construct validity and was the most preferred scale especially in younger and older pain patients with age less than 75 years. Chibnall and Tait (2001) found that the BS-21 had no error rate and high convergent validity in elderly without and with CI and had highest construct validity which did not have analysis regarding cognitive status because of the limitations of the sample size. In addition, the BS-21 emerged as the best scale regarding accuracy, convergent reliability, test-retest reliability, and construct validity, regardless of mental status and this study supported the use of the BS-21 in patients with mild to moderate CI (Chibnall & Tait).

The FPS is developed by Bieri et al. (1990) to measure pain intensity in children. It consists of seven-line-drawn faces presented in a horizontal format with different facial expressions that include a neutral face representing no pain, a severely

controlled face without tears representing worst pain, and five other facial expressions in between (McGuire et al., 2004). There were three reasons that the FPS was selected over other options. First, on this scale, the oval faces are more adult in appearance that makes the scale more acceptable to mature adults. Second, none of the faces has tears to avoid bias introduced by personal beliefs related to pain expression (McGuire et al.). Third, the use of neutral face to represent no pain rather than a happy face which could potentially add an emotional component of joy/happiness to the measure of pain intensity (Taylor & Herr, 2003). The participants were asked to select the face that best represented their pain intensity. In 2001, Hicks, von Baeyer, Spafford, van Korlaar, and Goodenough revised the seven-face FPS to the six-face FPS-R to make it possible to score on the widely used 0 - 5 or 0 - 10 metric. However, the seven-face FPS is considered more sensitive to represent pain intensity than the FPS-R which has six faces. Therefore, this study selected the FPS as one pain intensity scale. Many studies support the validity and reliability of the FPS. One study supported the construct validity and strong ordinal properties and strong test-retest reliability of the FPS for using in elderly without CI (Herr et al., 1998). The FPS could be comprehended by nondemented elderly, early AD patients, and midstage AD patients, 100%, 60%, and 30%, respectively (Scherder & Bouma, 2000). Taylor and Herr (2003) compared the FPS, the NRS, and the Iowa Pain Thermometer (IPT) in African American elderly without and with CI. They found that the FPS had strongest test-retest reliability in elderly with CI group; the FPS was the most preferred scale in both elderly without and

with CI groups which was not affected by gender or education; and convergent validity of the FPS was acceptable in the elderly with none to moderate CI. However, the FPS was also found having errors because some patients had less pain than the first face (Chibnall & Tait, 2001). In addition, the weak correlation was found between the FPS and other scales in the elderly with CI (.48 - .53) (Taylor et al., 2005).

2. Multidimensional assessment tools based on self-report.

The multidimensional assessment tools provide a comprehensive picture of pain. The two commonly used multidimensional pain assessment tools that based on self-report are the Short-Form McGill Pain Questionnaire (SF-MPQ) and Brief Pain Inventory (BPI).

In 1987, Melzack introduced a short form of the MPQ (SF-MPQ) which provided a shorter and quicker multidimensional measure of pain in clinical setting (McGuire et al., 2004). There are two sections in the SF-MPQ. The first section includes 11 sensory words (throbbing, shooting, stabbing, sharp, cramping, gnawing, hot-burning, aching, heavy, tender, and splitting) and four affective words (tiring-exhausting, sickening, fearful, and punishing-cruel). The second section consists of two measures of pain intensity. These two measures are the Present Pain Intensity Index and a 10 cm VAS with anchors of “*no pain*” and “*worst possible pain*”.

The Brief Pain Inventory (BPI) is a survey instrument originally for use

with cancer patients and can be used to measure pain in other conditions (McGuire et al., 2004). The BPI has been adopted in several countries for clinical pain assessment, epidemiological studies, and the studies of the effectiveness of pain treatment (Wang, Mendoza, Gao, & Cleeland, 1996). The BPI measures both the intensity of pain (Pain Severity Scale) and interference of pain (Pain Interference Scale). The Pain Severity Scale uses 0 - 10 numeric scales to measure the intensity of pain. The Pain Interference Scale assesses the interference of pain in the patient's life by using numeric scales. 0 is "no interference" and 10 is "interferences completely" and the patients are asked to rate the degree to which pain interferes with their general activity, mood, walking, normal work (includes both work outside the home and housework), relations with others, sleep, and enjoyment of life. The higher pain intensity means the higher pain interference (Cleeland, 1984). According to Serlin, Mendoza, Nakamura, Edwards, and Cleeland (1995), the pain interference scores of 1 - 4 might correspond as mild pain, scores of 5 - 6 as moderate pain, and scores of 7 or greater as severe pain. In 1996, Wang, Mendoza, Gao, and Cleeland developed a Chinese version of the Brief Pain Inventory (BPI-C) and demonstrated its reliability and validity in 147 cancer patients. The reliability of BPI-C was calculated separately. Coefficient alphas for four pain severity items (Pain Severity Scale) and the seven pain interference items (Pain Interference Scale) were .89 and .92, respectively. The validity of the BPI-C was assessed by confirmatory factor analysis. Factor analysis of the BPI-C items satisfied the criteria of reproducibility and interpretability. Therefore, the BPI meets common

standards of psychometric test development in the Chinese character format (BPI-C).

In this study, concurrent validity was assessed by examining the relationship between pain intensity and pain interference. Since the BPI-C was developed by Wang et al. (1996) and its high reliability and validity were demonstrated in that study, thus, the Pain Interference Scale as the subscale of the BPI-C was used as the present criterion to assess the concurrent validity of the five pain scales. In this study, the Pain Interference Scale was modified in order to fit with the subjects. The item “*normal work*” in the Pain Interference Scale was deleted because the participants in this study were all inpatients. The item “*ability to think and make decision*” was added as it was a common interference in patient with acute postoperative pain. Therefore, the modified Pain Interference Scale in this study had seven items to measure pain interference on general activity, mood, walking, relations with others, sleep, enjoyment of life, and ability to think and make decision. Subjects were asked to rate on 0 - 10 numeric scales with 0 equaling “*no interference*” and 10 equaling “*interfere completely*” regarding the above seven items.

Overall, the unidimensional pain assessment tools and the multidimensional pain assessment tools have their own advantages and disadvantages. The unidimensional pain assessment tools have become popular tools to quantify pain relief and pain intensity because they are easy to understand and use and place a minimal burden on the patients (Ho, Spence, & Murphy, 1996). In addition, the unidimensional pain assessment tools are more appropriate in assessing acute pain

rather than chronic pain as chronic pain is usually associated with other factors such as degree of support and depression (Ho et al., 1996). In contrast, the multidimensional pain assessment tools offer a comprehensive approach for assessing pain but the interpretation sometimes is difficult because of their complexity (Ho et al.).

Behavioral assessment tools. Behavioral assessment tools are based on the behavioral dimension of pain. The behavioral dimension of pain has two main components (McGuire et al., 2004). One component is the behaviors that are observable indicators of the presence and/or the severity of pain (e.g., grimacing, nonverbal vocalizations, communication with others, guarding and splinting, and fatigue). The other component is the behaviors that individuals use to relieve or control their pain (e.g., use of medications, positioning, and sleep/rest/activity patterns). There are some behavioral assessment tools in clinic such as the Post Anesthesia Care Unit (PACU) pain rating scale and the Checklist of Nonverbal Pain Indicators (CNPI).

Physiological assessment tools. Physiological assessment tools measure the physiological dimension of pain. Few physiological assessment tools are used in clinical setting because it is difficult to directly measure the etiology or organic origin of pain, levels of endogenous opioids and neurotransmitters, or other selected psychophysiologic factors (McGuire et al., 2004). However, the location, duration, and type of pain can be measured by instruments developed in recent years such as Pain

Map (McGuire et al.). In addition, the physiological indicators such as increased heart rate or blood pressure can be considered in the pain assessment process when self-report of pain is impossible in some patients.

In conclusion, there are mainly three types of pain assessment tools including pain assessment tools based on self-report, behavioral assessment tools, and physiological assessment tools. Behavioral indicators (e.g., restlessness, grimacing) and physiological indicators (e.g., increased heart rate, or blood pressure) can be considered in the pain assessment process when the patients can not report pain by themselves. However, it is also important to realize that the behavioral indicators or physiological indicators are not as reliable as the patient's self-report of pain. The behavioral and physiological indicators may be signs of other common conditions such as oxygen desaturation, stress, or anxiety, especially in the critically ill patients (Pasero, 2003) as well as in the postoperative patients. In order to assess pain correctly when using behavioral and physiological indicators, it is important to realize that an absence of these indicators should never be interpreted as the absence of pain. In addition, these indicators may be absent when the patient is experiencing severe pain or resolved before the pain is relieved (Pasero). In this study, the researcher focuses on the pain intensity scales which are under the category of the unidimensional pain assessment tools based on self-report. The reasons are the follows. Pain assessment tools based on self-report are the most reliable tools for assessing pain as pain is a subjective experience. In addition, among the pain assessment tools based on self-report, the

unidimensional pain assessment tools are more appropriate for assessing acute pain and they are easy to understand and use and place a minimal burden on the patient (Ho et al., 1996). Moreover, in the unidimensional pain assessment tools based on self-report, the pain intensity which is under the sensory dimension of pain is probably the most frequently assessed component of pain (Dahl, 1996; Gagliese & Melzack, 2006) because pain intensity is used to determine the effectiveness of pain management in clinic.

Psychometric Properties of Pain Intensity Scales in Adults

The most commonly used pain intensity scales are the VAS, the NRS, and the VDS in western countries. The VDS and the NRS may be appropriate across the adult lifespan according to many studies; for the VAS, though it has high validity, there are many limitations in the VAS especially when used in elderly patients. Gagliese et al. (2005) compared the VAS-horizontal (VAS-H), the VAS-vertical (VAS-V), the NRS, and the VDS in younger (less than 60 years old) and older (more than 60 years) surgical patients in the first 24 hours following surgery. They found that the NRS was the preferred scale and its properties were not age-related; the VDS was also good; however, the VAS was difficult to use among elderly including high rates of unscorable data and low face validity. Herr et al. (2004) compared the VAS, the NRS, the VDS, the Verbal Numeric Rating Scale (VNS), and the FPS among young and old adults by inducing an experimental painful stimulus. They found that all scales were

effective in discriminating different levels of pain sensation but the VDS was the most sensitive and reliable scale; the VAS had more failure rates and the scale most preferred in both groups was NRS followed by the VDS.

Besides the commonly used three scales, many studies found that the FPS and the BS-21 are suitable for pain intensity assessment in younger and older adults. Peters et al. (2007) compared the VAS-H, the VAS-V, the VDS, the Numerical Box-11 Scale (BS-11), and the BS-21 in younger and older chronic pain patients (divided into six age groups). They found that the number of mistakes on all scales increased with increasing age; however, the VAS had more mistakes. The BS-21 was the most preferred scale overall, even patients aged over 75 years especially preferred the VDS. Thus, the BS-21 is suitable for pain intensity assessment in heterogeneous groups and the VDS can be considered when the majority of patients are older adults (Peters et al.). Herr et al. (1998) evaluated the FPS for using with the community elderly and supported the validity and reliability of the FPS.

Many studies found that the elderly with CI are at an increased risk of experiencing pain (McGrath, Rosmus, Canfield, Campbell, & Hennigar, 1998; Weiner, Peterson, & Keefe, 1999). This may be due to the communication disorders, the characteristics of available pain measures, and the tendency of underreport pain in the CI patients (Stolee et al., 2005). In general, the elderly with CI received significantly less analgesics than the elderly without CI (Dawson, 1998). A study of nursing home residents found that 78% of the nursing home residents who had CI had a pain-causing

diagnosis, but fewer than 40% of them received analgesics (Feldt, Warne, & Ryden, 1998).

There are some studies comparing pain assessment tools in elderly patients with CI. The findings showed that patients with mild or moderate CI can use pain intensity scales to report pain reliably and validly but the qualities of pain scales are varied (Chibnall & Tait, 2001; Closs et al., 2004; Scherder & Bouma, 2000; Taylor et al., 2005; Taylor & Herr, 2003; Ware et al., 2006). Generally, the VDS, the NRS, the FPS, the BS-21, and the CAS were proved to be feasible for assessing pain in the patients with none to mild or moderate CI. The details would be presented under the topic of CI.

However, studies in pain assessment scale are limited in China. One study investigated the correlation between the NRS and the VDS in 50 Chinese patients after cardiac or general surgery (Zhao et al., 2002). They found that the NRS and the VDS (0-4) correlated well and suggested that a scale incorporating an NRS with a VDS was recommended. Another study by Li et al. (2007) compared the VAS, the NRS, the VDS, and the Faces Pain Scale Revised (FPS-R) in Chinese postoperative adults (18-78 years) from the first to the sixth postoperative days. They found that both the VDS and FPS-R had low error rates and nearly half of patients preferred the FPS-R but the VAS was difficult to understand and had highest error rates. However, the cognitive function was not measured in this study and the patients who could not complete the pain scales were excluded during the analysis. Therefore, considering the

lack of studies with Chinese population to guide the decision which pain scale might work best, this study would compare the psychometric properties of the evidence-supported tools: the VDS, the NRS, the FPS, the CAS, and the BS-21 in the adults varying in ages including the elderly with mild CI in Chinese population. Furthermore, the researcher also reviewed the study by Samabub, Petpichetchian, & Kitrungrrote (2009) which is similar to the present study but conducted in Thai population. Samabub et al. (2009) compared the VDS, the NRS, the FPS, the CAS, and the BS-21 and revealed that the pain intensity scale that had highest validity and reliability was the NRS for the young adult and middle-aged adult groups and the FPS for the elderly without and with mild CI groups.

Factors Influencing Psychometric Properties of Pain Scales

Based on the literature review, age, gender, educational level, cognitive impairment, and cultural background might influence the psychometric properties of pain scales.

Age

Age differences in the psychometric properties of pain scales are inconsistent from the literature review. Peters et al. (2007) compared pain intensity scales in younger and older patients and found that age was proved to be significantly related to making a mistake with older patients making more mistakes ($p = .02$) on the pain scales. In addition, patients of 75 years or older more preferred the VDS (p

= .014). In another study, Ware et al. (2006) compared pain scales in older minority adults and found the trends that elderly over 70 years preferred the FPS-R and elderly less than 70 years preferred the NRS. In contrast, Herr et al. (2004) compared pain scales in young and old adults and found that age did not impact failure to use pain scales, other conditions commonly associated with advanced age including cognitive and psychomotor impairment did since the cognitive and motor impairment were found to significantly increase the relative risk of failure to use the VAS successfully ($p < .05$). In addition, this study found that scale preference was not related to age. Gagliese and Katz (2003) compared measures of pain intensity and quality in younger and older surgical patients. They found that age differences in pain were dependent on the pain scale used; older patients had significantly lower scores than younger patients on the McGill Pain Questionnaire (MPQ) and Present Pain Intensity (PPI) but there were no differences on the VAS. In addition, several age differences in the psychometric properties of the scales were found in this study. On the 1st and 2nd postoperative days, the correlation between the VAS and the MPQ scores was significantly lower in the older than younger group; however, this may be due to that the MPQ might have inadvertently introduced fatigue effects especially for the older group. Gagliese et al. (2005) compared pain scales in younger and older surgical patients and found that increasing age was the only significant predictor of error rate since the patients who made errors on the VAS-horizontal (VAS-H) were older than those who did not make errors ($p \leq .009$). In this study, they also stated one unexplored

question that the error rate was influenced by the cognitive abilities or impacted by ageing was not clear as in this study only data from patients successfully completed each scale were included during analysis. In addition, they found that age was not related to scale preference. Another study Herr et al. (1998) evaluated the FPS for using with community elderly. In Herr et al.' study, the order of the faces presented was randomized and the subjects were asked to place the seven faces in order. This step was repeated after two weeks to determine stability of the rank ordering by the subjects. They found the tendency that the percentage of the elderly correctly ranking all the faces decreased with ageing on the first time. However, in this study the small number of subjects made age comparisons questionable. Li et al. (2007) compared pain intensity scales in Chinese postoperative adults and found that there were no significant differences in scale preference regarding age.

In summary, the findings that whether age influences the psychometric properties of pain scales are inconsistent. In terms of scale preference, Peters et al. (2007) found an age difference on scale preference, however, some other studies found that scale preference was not related to age (Gagliese et al., 2005; Herr et al., 2004; Li et al., 2007) and cognitive status (Herr et al.). For scale accuracy, some studies reported that age was significantly related to scale inaccurate responses (Gagliese et al.; Peters et al.). However, Herr et al. found that age did not impact failure to use the pain scale but cognitive and psychomotor impairment did. Moreover, overall, less study investigated the age differences in other aspects of psychometric properties.

Furthermore, since the development of physiology and psychology are different in the adult lifespan, this study would concern the age factor to determine whether age influenced the psychometric properties of the pain scales. The age groups in this study would be classified into three groups according to the changes of physiological and psychological development of the majority population in China. The standard classification of age group for research uses in China is: young adults (age 20 - 44 years), middle-aged adults (age 45 - 59 years), and old adults (age \geq 60 years) (Li et al., 2000; Yin et al., 2000).

Gender

Gender differences in the psychometric properties of pain scales are incongruent from the previous studies. Peters et al. (2007) compared pain intensity scales and found that female individuals more often preferred the VDS than males ($p = .041$). Li et al. (2007) found that there was a difference in the error rate regarding gender with more errors for the VAS among females ($p = .026$) but there were no significant differences in the scale preference regarding gender (Herr et al., 2004; Li et al.).

Educational Level

The findings regarding the educational level's difference in scale accuracy and preference were inconsistent from the previous studies. The result from

one study was unexpected that higher-educated patients made more mistakes in understanding of the ordinal properties of the scale ($p = .018$) (Peters et al., 2007). However, another study found that error rate was not related to educational level (Gagliese et al., 2005). In contrast, Li et al. (2007) found that there was a difference in the error rate between education levels with more errors for the NRS among those with more than high school educational level ($p = .022$) but no significant differences were found in scale preference regarding educational level. However, when Ware et al. (2006) examined the reliability and validity of the FPS-R, the VDS, the NRS, and the IPT in minority elderly, they found the trends that people with high school education or less preferred the FPS-R and people who had attended college or had earned degree preferred the NRS.

Cultural Background

Many studies found that cultural background influences the psychometric properties of pain scales. One study found that Chinese patients assessed pain more accurately with a vertical version of VAS than with the more commonly used horizontal version (Aun et al., 1986). The following two studies provide evidence that cultural background influences the construct validity of the FPS. The first study (Herr et al., 1998) evaluated the FPS for use with the European-American elderly. When they assessed the construct validity of the FPS that whether the elderly thought that the faces represented some level of pain, they asked the subjects to rate their

agreement on whether or not the faces represented six different constructs including pain, sourness, sadness, anger, boredom, and sleepiness. Based on the agreement of the subjects, they found that the FPS represented pain more strongly than any of other constructs but only significant difference between the construct of anger and all other concepts was found. Thus, faces could also represent other constructs except anger. However, when Taylor and Herr (2002) using the FPS with African-American older adults, they found that the construct of pain was significantly different from the constructs of sourness, sleepiness, and boredom, but no difference from sadness and anger. Moreover, the findings from the following three studies showed that the ethnic or culture differences might influence scale preference. Taylor and Herr (2003) compared the VDS, the NRS, the FPS, and the Iowa Pain Questionnaire (IPT) and reported that the FPS was the most preferred scale and had no failures in African-American old adults including the elderly with CI. In contrast, another two studies compared the similar scales in a primarily Caucasian sample, they found that the NRS and the VDS were the most preferred scales and had low errors in young and old adults without and with CI (Herr et al., 2004; Taylor et al., 2005), suggesting that the ethnic or culture differences might have an effect on scale preference. Therefore, the above evidences indicate that we need to consider cultural background when choosing and assessing pain assessment scales.

Overall, based on the review, age and CI are two main factors that might influence the psychometric properties of pain scales and these two factors would

be focused in this study. Since CI is an important concept in this study, the detail information about CI influencing the psychometric properties of pain scales and other information related to CI are presented under the following topic of CI.

Cognitive Impairment

Concept

Cognitive status consists of perception, memory, and thinking, including recognition/registration, storage, and use of information (McGuire et al., 2004). Cognitive impairment (CI) is one important aspect that influences the quality of life in the older adults (Petersen & Negash, 2008). There are many different kinds of CI such as mild CI and dementia. This part would focus on mild CI since the researcher studied the psychometric properties of pain assessment scales in the population including those with mild CI.

Mild CI is known as a transitional stage between normal ageing and early dementia (Li, Ng, Kua, & Ko, 2006; Petersen & Negash, 2008). Many studies are focusing the identification of mild CI because the therapeutic intervention in the early stage is more effective to arrest or reverse the cognitive decline (Li et al., 2006). It is generally accepted that there are two main subtypes of mild CI including amnesic mild CI and non-amnesic mild CI (Petersen & Negash). According to Petersen and Negash's review, for the amnesic mild CI, the memory deficits predominate and the

amnesic mild CI is also thought to represent the majority of person who will progress to Alzheimer's Disease (AD) over time; for the non-amnesic mild CI, there are no memory deficits but one or more of other cognitive domains (e.g., attention, execute function, visuospatial skills, and language) are impaired (Morris & Cummings, 2005; Petersen & Negash). In addition, the non-amnesic mild CI is considered to progress to non-Alzheimer's dementias such as frontotemporal dementia and vascular dementia (Albert & Blacker, 2006; Petersen & Negash).

Cognitive Impairment, Tool Selection, and Psychometric Properties of Pain Scales

The level of cognitive impairment (CI) influences the tool selection when assessing pain. Several studies showed that patients with mild or moderate CI can report pain reliably by using self-report pain scales (Chibnall & Tait, 2001; Closs et al., 2004; Scherder & Bouma, 2000; Taylor & Herr, 2003). However, pain assessment in patients with severe CI requires other strategies such as by using nonverbal behaviors, vocalizations, and family/caregiver reports to assess pain (Kovach, Weissman, Griffie, Matson, & Muchka, 1999). The level of CI also affects pain behavior and the psychometric properties of behavioral pain assessment tools. One study (Defrin et al., 2006) investigated whether the level of CI affects acute pain behavior and how it is manifested. They compared two pain behavioral tools including the Facial Action Coding System (FACS) and the Non-Communicating Children's

Pain Checklist (NCCPC) in patients with different levels of CI including mild, moderate, severe, and profound CI. They found that the patients with severe to profound CI exhibited high rates of “freezing reaction” (stillness) manifested mainly in the face during vaccination which was different from the behavioral response from the patients with mild to moderate CI. Thus, the FACS score did not show an elevation when it was used to assess pain in the patients with severe to profound CI. Therefore, this study suggested that the level of CI affects the acute pain behavior and pain assessment tool selection.

Some studies also found that the level of CI influences the psychometric properties of pain intensity scales. Closs et al. (2004) compared pain intensity scales among patients with different levels of CI and found that there was a tendency for less scales to be completed as the level of CI increased. They also found that all scales correlated significantly highly with each other in patients with none to moderate CI; however, the correlations between scales were far more variable at higher levels of CI and no scale correlated significantly with any another in the patients with severe CI. Another study (Chibnall & Tait, 2001) compared pain intensity scales in elderly without and with CI. They found that while reliability and validity of the pain scales were stronger in the patients with less CI, the levels found in the patients with more CI were also acceptable; however, the largest discrepancy related to mental status was the test-retest reliability. The test-retest reliability in the patients with CI, only when restricted to a 3-day period, approximated the reliability obtained in the patients

without CI for a 7-day period. Therefore, they suggested that the temporal stability of ratings begins to decrease for patients with CI after about three days. Furthermore, the study by Taylor and Herr (2003) compared pain intensity scales in African American elderly without and with CI groups. They found that test-retest reliability coefficients at a 2-week interval were acceptable and ranged from .52 to .83 in two groups but the weaker relationships were found in the elderly with CI group which might be related to the memory impairment in this group. In contrast, Taylor et al. (2005) examined the VDS, the NRS, the FPS, and the IPT in the elderly with CI group and found that the test-retest reliability over a 2-week interval was unacceptable ($\rho = .26 - .67$). Furthermore, Scherder and Bouma (2000) compared pain intensity scales between an early and mid-stage of AD population and nondemented elderly population. They found that the ability to comprehend the pain scales was decreased with the increased level of CI.

In summary, the level of CI not only influences the pain behaviors and the selection of pain scales when assessing pain in patients with different levels of CI but also influences the psychometric properties of pain assessment scales especially the test-retest reliability.

Psychometric Properties of Pain Scales in Elderly with CI

There are some studies comparing pain assessment tools in elderly patients with CI. The findings showed that patients with mild or moderate CI can use

pain intensity scales to report pain reliably and validly but the qualities of pain scales are varied. Taylor et al. (2005) evaluated the VDS, the NRS, the FPS, and the IPT in Caucasian old adults without and with CI. They found that the concurrent validity of these scales was good and both the elderly without and with CI groups were able to use these pain scales and preferred the IPT and the VDS, however, the test-retest reliability at a 2-week interval was unacceptable for most scales in the elderly with CI group and suggested that the stability issue in the elderly with CI must be considered in future studies. Ware et al. (2006) investigated the VDS, the FPS-R, and the IPT in a majority of African American old adults without and with CI. They supported the validity and reliability of these scales and revealed that the NRS was the preferred scale in the elderly without CI group and the FPS-R was the preferred scale in elderly with CI group. In addition, when race was considered, African Americans preferred the FPS-R. In contrast, Chibnall and Tait (2001) compared the VDS, the FPS, the horizontal 21-point (0-100) box scale (BS-21), and the vertical 21-point (0-20) box scale in elderly without and with CI in a subacute care facility over a 14-day period. They found that the BS-21 was suitable for pain assessment in older patients including those with mild to moderate CI and the elderly with CI could rate pain reliably and validly. Differently, Taylor and Herr (2003) compared the FPS, the NRS, and the IPT in African American elderly without and with CI. They found that both elderly without and with CI groups preferred the FPS to represent the intensity of their pain and all scales were easy to use and understandable. Closs et al. (2004) compared the VDS, the

NRS, the FPS, the CAS, and the Mechanical Visual Analog Scale (MVAS) among elderly with different levels of CI. They found that the VDS was the most easily understandable scale and appeared to be suitable for all but the patients with most severe CI and the NRS was also good following the VDS. The CAS is another tool that might be feasible in using with the elderly with CI. Scherder and Bouma (2000) compared the CAS, the FPS, and the Facial Affective Scale (FAS) among an early and midstage of Alzheimer's disease (AD) population and nondemented elderly population. They found that the CAS could be comprehended very well in nondemented elderly patients and the early AD patients. In sum, based on the above evidences, the VDS, the NRS, the FPS, the BS-21, and the CAS can be attempted to assess pain in the patients with none to mild or moderate CI.

Cognitive Function Assessment

Since the patients with mild CI would be included in this study, the researcher reviewed the instruments for screening cognitive status. The Mini-Mental State Examination: MMSE (Folstein, Folstein, & McHugh, 1975) is a global instrument for screening CI or dementia in both clinical settings and community settings (Commenges et al., 1992; Kukull et al., 1994). There are seven dimensions which cover a broad range of cognitive domains in this scale. These seven dimensions include orientation to time (5 points), orientation to place (5 points), registration of three words (3 points), language (8 points), recall of three objects (3 points), attention

and calculation (5 points), and visual construction (1 point). In addition, the MMSE consists of 30 items of dichotomous questions. The total possible scores range from 0 to 30 points. The cut-off point is 24. A score of 23 or less indicates the presence of CI (Tombaugh & McIntyre, 1992).

The MMSE has been translated into Chinese, however, the researcher needs to consider sociocultural differences when adapting tests and using clinically (Okamoto, Case, Bleiker, & Henderson, 1996). There are mainly three types of the Chinese version of MMSE according to the literature review.

The first type of the Chinese version of MMSE was the CMMSE that was minimally modified culturally and formally to keep it in accordance with the original MMSE (Yu et al., 1989). Zhang (1991) compared several screening instruments for dementia and found that the CMMSE was better than other instruments with sensitivity 92.5% and exceptionality 79.1% for screening dementia. This CMMSE was used widely and many epidemiological studies used it in dementia screening (Zhang et al., 1999; Zhou, Furgang, & Zhang, 2006). However, there are many different versions of the cutoff points and it is important to choose a high valid one. Based on the literature review, the cutoff points in Zhang et al. (1999) are more rationale. In order to offer a benchmark for cutoff points, Zhang et al. examined the distribution of the CMMSE scores in terms of ages and educational levels in the Chinese residents aged 55 years and over. The residents in Zhang et al.'s study were recruited from the urban and rural areas of Beijing. They found the bivariate

correlation between the CMMSE scores and all three factors of age, sex, and educational level ($p < .01$). The optimal cutoff values were taken from the 10th percent lowest score in the age group of 60 to 65 years for each educational level: illiterate ≤ 19 , elementary school ≤ 22 , and secondary school or higher ≤ 26 . When compared with the cutoff points used by other studies, these values achieved higher sensitivity with 90.7% in the urban area and 97.1% in the rural area. Therefore, these cutoff points would improve the disease detection by reducing the number of false negatives. The identified cutoff points in Zhang et al.'s study were used in many dementia screening studies (Zhang et al., 2005; Yang et al., 2008). Yang et al. investigated the prevalence and risk factors of CI among the elderly in China. They found that 14.14% of the elderly had mild CI and 5.17% of the elderly had severe CI and CI was related to the grade of education, age, and gender. In Yang et al.'s study, the elderly without CI group included illiterate people with the CMMSE score > 19 , people having primary school educational level with score > 22 , and people having secondary school educational level or higher with score > 26 . The cutoff points for mild CI group were 17-19 for illiterate people, 20-22 for people with primary school educational level, and 24-26 for people with secondary school educational level or higher.

The second type of the Chinese version of the MMSE is the Chinese adopted Mini-Mental State Examination (CAMSE). As illiteracy is prevalent among the current elderly Chinese, Xu et al. (2003) developed this CAMSE for screening dementia among illiterate or less educated elderly Chinese. In this CAMSE, the

literacy-dependent items of the MMSE were modified or substituted by equivalent items that were not literacy-dependent. Some items were modified according to the sociocultural compatibility. The main structures of the test were kept intact with the MMSE and similar principles for scoring were used. Contents of the test items in the CAMSE were kept in accordance as much as possible with the original MMSE so that CAMSE could test the same cognitive functions. Changes were made with reference to other Chinese translations of the MMSE. Performing CAMSE takes about 15 minutes. After developing it, the CAMSE was administered to 370 elderly outpatients. Sensitivities and specificities for detecting dementia were evaluated by adjusting for different CAMSE cut-off points. The principle that higher sensitivities with acceptable specificities were preferred was used to decide the cutoff points. The optimal cutoff points of 22 for literates and 20 for illiterates yielded a sensitivity of 83.87% and a specificity of 84.48%. Corresponding positive predictive value (PPV) was .65 and negative predictive value (NPV) was .94. Illiterate subjects got a higher total score than literate subjects ($p < .05$). Test-retest reliability was .75 ($p < .01$). In this study, severe dementia patients were excluded which ensured that the optimal cutoff values as well as their sensitivities and specificities would be particularly relevant for the detection of mild dementia. This study recommended that the CAMSE was feasible for use in clinical settings for dementia screening.

The third type of the Chinese version of the MMSE is the 16-item CMMSE. According to some studies (Lou, Dai, Huang, & Yu, 2003; Smith, Breitbart,

& Platt, 1995), every item in the MMSE is not equally efficient at identifying cognitive disturbance and certain items are more sensitive to detect the changes than other items. In order to identify the most efficient items from the MMSE, Lou, Dai, Huang, and Yu (2007) developed the 16-item CMMSE. In their study, they used the item response theory to identify the most efficient items from the MMSE for cognitive function assessment. The identified 16 items were mainly related to the measures of orientation, recall, attention, and calculation and these identified items had some consistency compared with the previous studies. The internal consistency of the 16 items was .84. The proposed new cutoff point of 16-item MMSE was 11. The score of 11 was determined for the purpose of over-identifying the patients who were at risk so as to ensure early detection of the onset of cognitive disturbance. When compared with the original 30-item MMSE from the cutoff point of 24, the correct classification rate of the 16-item CMMSE was .94, the sensitivity was 100%, and the specificity was 97.4%. Therefore, this study showed that a few items were needed to describe the subject's cognitive status and the MMSE could be simplified. Deleting the items with less variation makes the tool shorter, easier to administer, and less strenuous for the participants, but also maintains validity.

In conclusion, for the CAMSE, it was good for screening dementia among illiterate or less educated Chinese elderly but the cutoff points for screening mild CI patients were not identified. For the 16-item CMMSE, it was shorter and easier to administer and also had validity but the identified cutoff point was to detect the

onset of cognitive disturbance and the cutoff points for screening mild CI were also not identified. For the CMMSE, the cutoff points in Zhang et al. (1999) achieved high sensitivity and these cutoff points were also adapted in accordance with the educational level which is an important factor that influences the CMMSE score. Therefore, the cutoff points of the CMMSE from the previous studies (Yang et al., 2008; Zhang et al.) would be used to divide the elderly into elderly without CI group and elderly with mild CI group in this study. Finally, the elderly without CI group included illiterate people with score > 19, people having primary school education with score > 22, and people having secondary school education or higher with score > 26. The elderly with mild CI group included illiterate people with score 17-19, people having primary school education with score 20-22, and people having secondary school education or higher with score 24-26.

Treatment for the Mild Cognitive Impairment

For the treatment of mild CI, many clinical trials on mild CI are being undertaken (Petersen & Negash, 2008). Donepezil, cholinesterase inhibitors, galantamine, and rivastigmine have been investigated in mild CI. However, there are no pharmacologic interventions demonstrated to be efficacious for treating mild CI (Petersen & Negash). Although there are no effective drugs for treating mild CI, the no drug treatments have been approved by the Food and Drug Administration (FDA) for the indication of mild CI (Morris & Cummings, 2005). The management of mild CI is

currently non specific. The proposed methods may involve control of vascular risk factors, treatment of a concomitant condition such as depression or hypothyroidism, and reduction in the use of anticholinergic drugs (Mariani, Monastero, & Mecocci, 2007). Besides the above approaches for treating mild CI, the psychosocial and nutritional interventions for preventing cognitive decline have also gained considerable attention recently (Petersen & Negash). For example, several studies have found that frequent participation in cognitively stimulating activities can protect against cognitive decline and reduce the progress to AD (Petersen & Negash). These cognitively stimulating activities can involve reading a book, playing a game, or listening to a radio program. In addition, an active and socially integrated lifestyle in late life protects against dementia and AD (Mariani et al., 2007). The nutrition intervention is also found to have the ability for protecting against cognitive decline (Petersen & Negash). These nutrition interventions may involve low consumption of total fats, saturated fatty acids, and cholesterol.

Impacts of Surgery on Cognitive Function

Surgery is a stressful event for older patients and alters their cognitive function (Lou et al., 2003). It is generally accepted that acute confusional state (ACS) is an indicator of cognitive disturbance and is common among hospitalized postoperative patients especially in elderly (Duppils & Wikblad, 2000; Rudberg, Pompei, Foreman, Ross, & Cassel, 1997). The clinical features of the ACS may

include reduced ability to focus, shift attention, and thinking and speech disorganization (Lou et al.). Many studies investigated the cognitive changes in the patients after surgery. One study (Duppils & Wikblad) investigated the ACS in 225 elderly patients undergoing hip surgery and the MMSE was used to measure cognitive function. They found that the incidence of the ACS was 20% among non-confused hip-surgery elderly patients aged 65 or more and those patients with mild or moderate CI more often developed the ACS. This study also found that the onset of the ACS was 3 - 45 hours after surgery in all but eight patients and the duration of the ACS among recovered patients was generally less than 48 hours. Another study assessed the postoperative cognitive changes among 106 older Taiwanese patients (Lou et al.) and the cognitive function was assessed four times (admission, onset of ACS, ACS day 3, and ACS day 5) by using the MMSE. They found that the subjects who experienced the ACS had significantly lower MMSE scores than non-ACS patients and the total scores on the MMSE decreased from time 1 assessment to time 2, but increased again through time 3 and time 4.

This study would assess the psychometric properties of pain scales in the postoperative elderly patients with mild CI. Based on the above studies, the researcher would start to assess pain intensity from 49 hours after surgery to avoid the influence of cognitive disturbance that caused by surgery.

Summary

The literature review starts with the major concepts of this study pain in postoperative patients. Pain is a subjective feeling and also a multidimensional experience. Two theory including gate control theory and nociception theory are essential for understanding pain or postoperative pain mechanisms. Postoperative pain is the most common type of acute pain. The unrelieved postoperative pain not only causes negative physiological effects and psychological problems but also leads to socioeconomic issues. Elderly tend to develop the above problems because of their increased vulnerability to stressors. Therefore, postoperative pain management is so important that we need to pay attention to especially in elderly. It is well established that accurate pain assessment is a prerequisite for effective pain management. To achieve accurate pain assessment, the instruments that used for assessing pain should be quantified. To quantify a subjective construct, such as pain, the psychometric properties of its instrument including validity and reliability need to be assessed.

Therefore, the researcher reviews the measurement theory and psychometric properties of the instrument including validity and reliability. Based on the feasibility, validity would be evaluated by face validity, concurrent validity, and convergent validity and reliability would be evaluated by test-retest method in this study.

After that, the types of pain assessment tools are reviewed. Based on the review, there are mainly three types of pain assessment tools including pain assessment

tools based on self-report, behavioral assessment tools, and physiological assessment tools. However, pain assessment tools based on self-report are the most reliable tools for assessing pain as pain is a subjective experience. Moreover, pain intensity is probably the most frequently assessed component of pain. Therefore, the researcher reviews the psychometric properties of pain intensity scales that based on self-report in adults. The findings of western research showed that the qualities of pain assessment scales use in various age groups including elderly patients with CI are varied. Overall, the BS-21, the FPS, the CAS, the NRS, and the VDS can be attempted to assess pain in the patients with none to mild or moderate CI.

Then, factors influencing pain measures are reviewed. Among these factors, age and CI which are two main factors that might influence the psychometric properties of pain scales would be two focuses in this study. Thus, the researcher reviews the relevant information about age and CI. Moreover, some studies found that culture also influences the psychometric properties of pain scales. Considering the culture difference, however, studies in pain assessment scales are limited in China. In addition, one recent Chinese study was conducted but the cognitive function was not measured and the patients who could not complete the pain scales were excluded during the analysis.

In clinical setting, it is important to identify the high valid and reliable scales for assessing pain across the adult lifespan. Although a few studies are from Chinese adults with surgery, there is a lack of studies to compare pain assessment

scales in various age groups: postoperative adult patients, elderly patients without CI, and elderly with mild CI. As the number of elderly patients with surgery has increased in China, this study would examine the psychometric properties of the evidence-supported tools: the VDS, the NRS, the BS-21, the FPS, and the CAS in the adults varying in ages including the elderly with mild CI in Chinese population.

CHAPTER 3

RESEARCH METHODOLOGY

In this chapter, the methodological issues including research design, research setting, population and sample, instrumentation, ethical considerations, and data collection procedures are presented as follows.

Research Design

This was a descriptive comparative study among four groups of subjects: young adults (age 20 - 44 years), middle-aged adults (age 45 - 59 years), elderly (age \geq 60 years) without CI, and elderly (age \geq 60 years) with mild CI. This study was conducted in a teaching hospital in Kunming, China.

Research Setting

This study was conducted in the Second Affiliated Hospital of Kunming Medical University which is located in the southwest of China. This hospital is one of the biggest teaching hospitals in Kunming.

In this hospital, the total number of surgical cases was 12,270 in 2007 and increased to 13,364 in 2008 (The Second Affiliated Hospital of Kunming Medical College Statistics, 2007, 2008). Based on the number of patients in the wards, the surgical wards selected in this study were general surgical ward,

hepatic-bilious-pancreatic surgical ward, abdominal microsurgery ward, and urology surgical wards.

General Information of Preoperative Procedures

The surgical cases are referred from the surgical outpatient department routinely. After being confirmed for surgery, patients are scheduled for operation. Then the patients are admitted on the surgical wards commonly two days before the scheduled operation date. During the preoperative period, patient's consent form for a surgical procedure and investigation are taken. The doctors and nurses will prepare the patients for operation. The anesthesiologists also evaluate the patient's health condition before operation. Routinely, the general surgeries are operated from Monday to Friday every week.

Information for Postoperative Pain Management

Generally, the patients are transferred to their original wards from the Post Anesthetic Care Unit (PACU) when their conditions are stable. However, the patients whose conditions are serious or unstable are transferred to the Surgical Intensive Care Unit (SICU) after surgery. There are some similarities and differences in postoperative pain management between the SICU and general surgical wards. The similarities of postoperative pain management in SICU and general surgical wards are the following three aspects. Firstly, the present techniques for postoperative pain

management in SICU and general surgical wards are Patient Controlled Analgesia (PCA), Patient Controlled Epidural Analgesia (PCEA), lumbar epidural cocktail, continuous IV, and oral medications. Secondly, the common analgesics prescribed for pain relief are opioids (morphine, fentanyl, and pethidine) and local anesthetics. The oral medications usually prescribed are the non-steroidal anti-inflammatory drugs (NSAIDs). Thirdly, the postoperative pain is managed by the doctors until the patient is discharged. However, there are also some differences in the postoperative pain management especially in the postoperative pain assessment between the general surgical wards and SICU. For one thing, postoperative pain is assessed by using the Prince-Henry Scale in SICU. The Prince-Henry Scale consists of five numerically ranked descriptors. 0 means that the patients feel no pain when coughing; 1 means that the patients can feel pain when coughing; 2 means that the patients can feel pain even when breathing deeply; 3 means that the patients can feel mild or moderate pain when at rest; and 4 means that the patients can feel severe pain when at rest and can not bear the pain. Therefore, the Prince-Henry scale is mainly used for assessing pain interference. The patients in the SICU usually have large abdominal operation which may interfere with their breathing. In order to avoid the respiratory complications caused by postoperative pain, the Prince-Henry Scale is appropriate for monitoring the pain interference in the SICU. Compared with the SICU, pain intensity is commonly used to determine the effectiveness of pain management in the general surgical wards; however, pain intensity is assessed by the nurse without objective pain assessment

tools. For another, postoperative pain management is more accurate in the SICU than in the general surgical wards.

In this study, the researcher assessed the psychometric properties of the five pain intensity scales which are different from the pain interference measurement Prince-Henry Scale. In order to avoid the disturbance of the Prince-Henry Scale to the patients, this study was conducted in the general surgical wards and the patients who stay in the SICU were excluded.

Population and Sample

Population

The target population in this study was postoperative adults. As age and CI were two focuses of this study, the age group and CI level needed to be categorized. The standard classification of age group for research use in China is: young adults (age 20 – 44 years), middle-aged adults (age 45 – 59 years), and old adults (age \geq 60 years) (Li et al., 2000; Yin et al., 2002). In addition, the cutoff points of the Chinese version of Mini-Mental State Examination (CMMSE) were used to divide the elderly into elderly without CI group and elderly with mild CI group. (Zhang, 1998). Therefore, this study would divide the postoperative adults into four groups: young adults (age 20 - 44 years), middle-aged adults (age 45 – 59 years), elderly (age \geq 60 years) without CI, and elderly (age \geq 60 years) with mild CI. The elderly without CI group included illiterate

people with score > 19, people having primary school education with score > 22, and people having secondary school education or higher with score > 26. The elderly with mild CI group included illiterate people with score 17-19, people having primary school education with score 20-22, and people having secondary school education or higher with score 24-26.

Samples Size Estimation

The proportional estimation from the population was used to determine the sample size in this study. The average total number of surgical cases in the hospital was about 12,817 during the years of 2007 and 2008 (The Second Affiliated Hospital of Kunming Medical College Statistics, 2007, 2008). A sample of 1% of the population for the larger population (>10,000) is considered suitable for descriptive studies (Singchangchai, Khampalikit, & Na-sae, 1996). Therefore, by taking 1% of the total number of surgical patients, the sample size in this study was 128. As a larger sample size is more representative of the population and subgroups analysis would be also conducted in this study, the sample size was increased to 200 of which 50 patients were needed to comprise each group (young adults, middle-aged adults, elderly without CI, and elderly with mild CI).

Inclusion Criteria

Subjects in this study were recruited by using purposive sampling from

the selected surgical wards. The inclusion criteria were: age over 20 years, admission for scheduled operation, no more than a mild CI level for elderly aged ≥ 60 years according to the CMMSE (score ≥ 17 if illiterate, ≥ 20 for people with primary school educational level, ≥ 24 for people \geq secondary school educational level; Yang et al., 2008), good eyesight including those using corrective lenses but excluding color blinded patients, good hearing including those using hearing devices, able to communicate in Mandarin, and willing and able to participate.

Instrumentation

The instruments used in this study consisted of the instruments for collecting background data, the instruments for assessing patient's pain intensity, and the instruments for assessing the validity of pain intensity scales. The instruments for collecting background data were the Demographic Data Questionnaire (DDQ) and the CMMSE. Instruments for assessing patient's pain intensity were the Patient's Pain Intensity Assessment Form One and Patient's Pain Intensity Assessment Form Two. The instruments for assessing the validity of the pain intensity scales included the Scale Preference Questionnaire (SPQ), the Scale Simplicity Questionnaire (SSQ), the Scale Accuracy Checklist (SAC), and the modified Pain Interference Scale (PIS). The descriptions of each instrument and the validity and reliability of the instruments are presented as follows.

Instruments for Collecting Background Data

Demographic Data Questionnaire (DDQ) (Appendix B1)

DDQ was designed by the researcher to collect general demographic and socioeconomic characteristics and health related information. This questionnaire consisted of two sections. The first section consisted of the questions on general demographic and socioeconomic characteristics including questions on age, gender, religion, educational level, occupation, and marital status. The second section contained the questions on general health data including date of surgery, operation sites, postoperative pain medications, and previous history of surgery. The DDQ was completed by the researcher according to the patient's medical records or by interview.

Instrument for Assessing Cognitive Function: Chinese Version of Mini-Mental State Examination (CMMSE) (Appendix B2)

The Mini-Mental State Examination (MMSE) is a widely used and well-validated instrument for screening cognitive function (Kukull et al., 1994; Tombaugh & McIntyre, 1992). A Chinese version of MMSE (CMMSE) was used in this study with possible scores ranging from 0 to 30. Since the CMMSE scores have been found to be related to the grade of education, the cutoff points of the CMMSE, which were demonstrated high sensitivity in differentiation CI regarding different educational levels in the prior studies (Zhang et al., 1999; Yang et al., 2008), were used

in this study. The CMMSE cutoff points divided the elderly into two groups: without CI group and with mild CI group. The elderly without CI group included illiterate people with score > 19, people having primary school education with score > 22, and people having secondary school education or higher with score > 26. The elderly with mild CI group included illiterate people with score 17-19, people having primary school education with score 20-22, and people having secondary school education or higher with score 24-26.

Instruments for Assessing Patients' Pain Intensity

The Patient's Pain Intensity Assessment Form One (Appendix B3) and the Patient's Pain Intensity Assessment Form Two (Appendix B4) were used for assessing the pain intensity of patients. On each form, pain intensity was measured by the five pain intensity scales including the NRS, the VDS, the FPS, the CAS, and the BS-21 as mentioned in chapter two.

On the Patient's Pain Intensity Assessment Form One, all the five pain scales and their corresponding instructions were presented. The patients were asked to rate the intensity of their vividly remembered painful experience by using these five pain scales.

On the Patient's Pain Intensity Assessment Form Two, each pain scale and its corresponding instruction were presented on a single sheet. In addition, on each sheet, four different pain intensity ratings, including current pain score and daily

retrospective worst, least, and average pain scores, were assessed. Therefore, this form consisted of five sheets. To avoid the order effects, each type of the scale should have the chance to occur first and occur last (Peters et al., 2007). Thus, five different versions of this form were randomly selected from all the possible versions. This study used the following five versions: (1) VDS, NRS, CAS, FPS, BS-21; (2) NRS, VDS, BS-21, CAS, FPS; (3) CAS, FPS, VDS, BS-21, NRS; (4) FPS, BS-21, CAS, NRS, VDS; and (5) BS-21, FPS, VDS, NRS, CAS. Moreover, each patient could randomly receive one of these five different versions of this form.

In addition, for the above two forms, to avoid the influence of impaired visual abilities accompanying with ageing or drug use, all pain scales were consistently presented in a large format (14-point). For the FPS, as used in the previous study (Taylor & Herr, 2002), the height of the faces was increased to 4 cm and the facial markings were darkened to enhance visualization of facial characteristics.

Instruments for Assessing the Validity of the Pain Intensity Scales

The Scale Preference Questionnaire (SPQ), the Scale Simplicity Questionnaire (SSQ), and the Scale Accuracy Checklist (SAC)

In this study, face validity of the five pain intensity scales was assessed by three aspects regarding scale preference, simplicity, and accuracy. These three aspects were assessed separately by the following three instruments.

Firstly, the SPQ (Appendix B6) was used to assess scale preference. On the SPQ, five pain intensity scales were presented and the participants were asked to rank order from 1 to 5 when 1 = most preferred and 5 = least preferred. Secondly, the SSQ (Appendix B7) was used to assess scale simplicity. On the SSQ, five pain intensity scales were presented and the participants were asked to rank order from 1 to 5 when 1 = simplest and 5 = least simple. Lastly, scale accuracy was evaluated by the number of the subjects with accurate response (without any errors) and with inaccurate response (with any errors) by using the SAC (Appendix B8) as mentioned in Chapter 2 (p. 39). In conclusion, for assessing face validity, the higher number of subjects indicating the most preferred scale and the simplest scale, and the higher number of the subjects with accurate response of the scale indicated the higher face validity.

The Modified Pain Interference Scale (PIS) (Appendix B5)

Concurrent validity of the pain intensity scales was assessed by examining the relationship between the pain intensity and pain interference. Thus, the modified PIS as the subscale of the BPI-C was used as the present criterion to assess the concurrent validity. The modified PIS in this study had seven items to measure postoperative pain interference on general activity, mood, walking, relations with others, sleep, enjoyment of life, and ability to think and make decision. Subjects were asked to rate on 0 - 10 numeric scales with 0 equaling “no interference” and 10 equaling “interfere completely” regarding the above seven items.

Validity and Reliability of the Instruments

Assessing Content Validity of the Scale Accuracy Checklist (SAC) and the Modified Pain Interference Scale (PIS)

The SAC for assessing face validity and the modified PIS for assessing concurrent validity were analyzed for content validity by a panel of three Thai experts consisting of two nursing experts in research methodology and one pain expert from anesthesiology department. They were asked to independently rate the relevance of each item with regard to its objective whether it measures what it should be measured by using a 4-point rating scale: (i) not relevant, (ii) somewhat relevant; (iii) quite relevant; and (iv) very relevant. Finally, the Content Validity Index (CVI) was computed equaling the proportion of items giving a rating of quite relevant and very relevant by those three experts. Finally, the CVI was 1 for the SAC and 1 for the modified PIS. No change was made from the experts' feedback.

Assessing Reliability of the Modified Pain Interference Scale (PIS)

The internal consistency reliability of the modified PIS was analyzed by using Cronbach's alpha. Twenty patients including five patients for each group (young adults, middle-aged adults, elderly without CI, and elderly with mild CI) were recruited for assessing the internal consistency of the modified PIS in the pilot study. Finally, the internal consistency reliability of the modified PIS was proved to be very

good since the Cronbach's alpha was .84 from the pilot testing (N = 20) and .90 from the actual study testing (N = 200).

Ethical Considerations

1. This study was approved by the Institutional Review Board (IRB) of Faculty of Nursing, Prince of Songkla University, Thailand, and was also granted permission by the Second Affiliated Hospital of Kunming Medical University, China.

2. The nurses in the ward introduced the researcher to the patients when there were potential subjects in the ward. The potential subjects were told the purpose of the study and how they would be involved. They were also informed that they had the right to stop or discontinue the interviewing process based on their own reasons without fear of any negative consequence to the care provided to them during their hospitalization. Subject confidentiality was also maintained.

3. A consent form (Appendix A) was given based on the subject's interest. Signed informed consent or patient's verbalization of willingness to participate was used as the sign of their consent.

4. There was no evidence shown risk related to completing the questionnaires. However, there is a possibility that self-rating of the pain level would make the patients feel physical or psychological discomfort or the questionnaires used in this study would be a burden to the patients, disturbing their rest after surgery. During the data collection, approximately 20 subjects reported tiredness, the researcher

stopped the interview and let the patients take a rest and helped the patients to solve the discomfort if they needed. Then, the researcher continued the interview when the patients felt better and were willing to continue.

Data Collection Procedures

Preparation Phase

The researcher obtained approval from the Institutional Review Board of Faculty of Nursing, Prince of Songkla University, and the Second Affiliated Hospital of Kunming Medical University. Once permission was granted, a letter was sent to the selected surgical wards. A pilot study was conducted on 20 surgical patients including five patients for each of the four groups. The purposes were to determine the feasibility of the proposed study, to check the reliability of the instrument, and to identify any problems encountered in the data collection process. The patients recruited in this pilot study were excluded in the sample of the study.

Data Collection Phase

- 1) Operation schedule sheets were reviewed from Sunday to Thursday (5 days/week) by the researcher.
- 2) On the preoperative day, the researcher reviewed the patients' records and interviewed the patients. Those who met the inclusion criteria were

approached to determine their interest in participation. Signed informed consent or verbalization of willingness to participate was obtained and the Demographic Data Questionnaire was completed by the researcher according to the medical records or by interview. For the elderly aged ≥ 60 years, the CMMSE was administered to them and the elderly who had more than a mild CI level according to the CMMSE were excluded in the study. Then, subjects were shown the pain scales and explained how to use them and the Patient's Pain Intensity Assessment Form One was given to each patient to assess the vividly remembered painful experience.

3) During 49-72 hours after surgery, the Patient's Pain Intensity Assessment Form Two was administered to each patient. The patients were asked to rate their current operative pain intensity and retrospective worst, least, and average pain intensity during the past 24 hours by using the five pain scales. The patients completed the scales by marking, or pointing, or verbally stating. For this form, each patient could randomly receive one version from the five different versions. In order to treat the patients equally and make the form distribution easy to manage, the patients in each group from number 1 to number 10 received version one, 11-20 received version two, 21-30 received version three, 31-40 received version four, and 41-50 received version five (Figure 2). After completion, the modified Pain Interference Scale (PIS) was given to them and they were asked to rate the degree to which the worst pain during the past 24 hours had interfered with them regarding the seven items in the modified PIS.

4) During 73-96 hours after surgery, the Patient's Pain Intensity Assessment Form One was given to each patient and the vividly remembered painful experience was assessed again. At last, the Scale Preference Questionnaire and Scale Simplicity Questionnaire were administered to the patients.

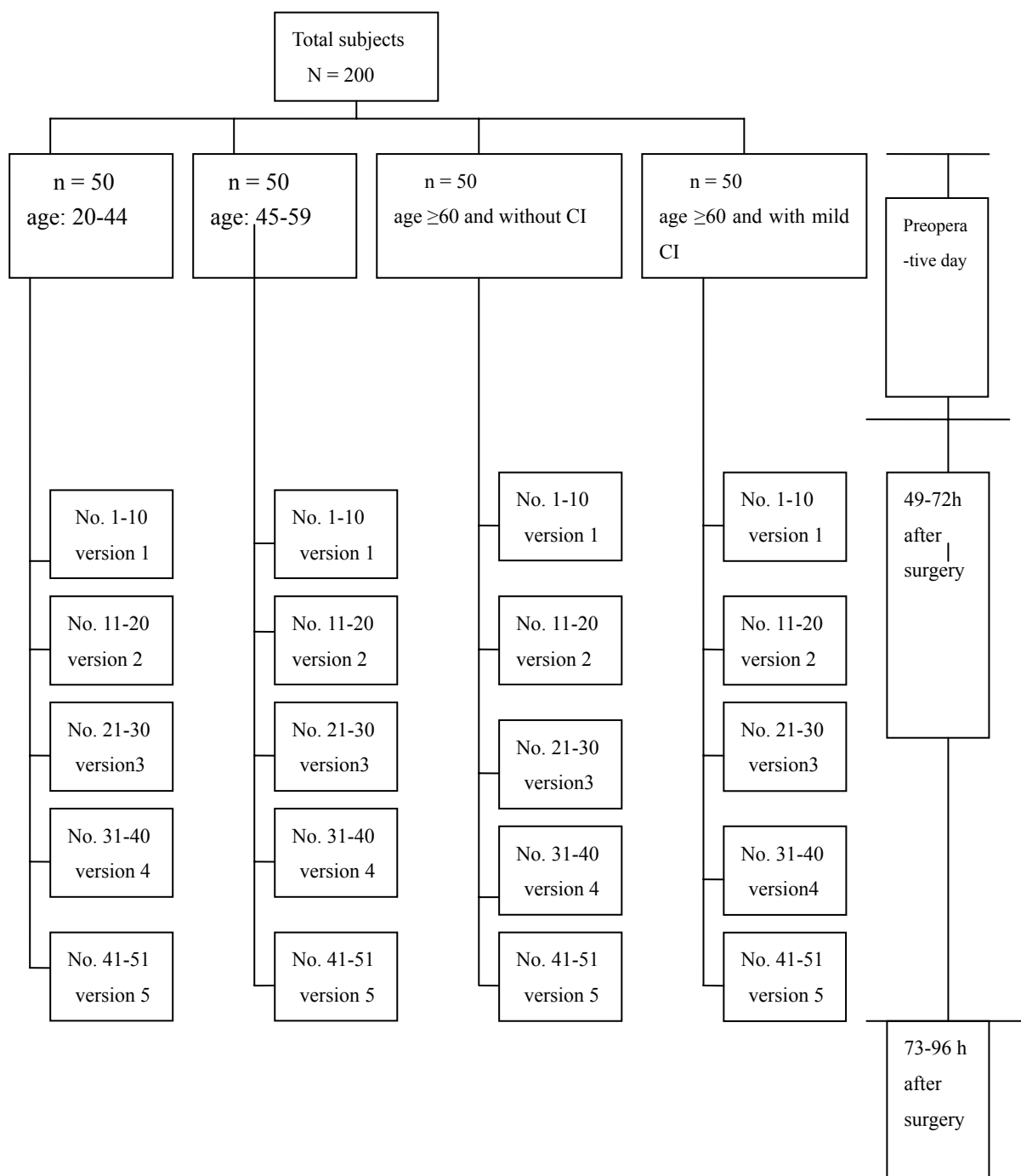


Figure 2. Data collection procedure

Note. Version 1 = VDS, NRS, CAS, FPS, BS-21; Version 2 = NRS, VDS, BS-21, CAS, FPS; Version 3 = CAS, FPS, VDS, BS-21, NRS; Version 4 = FPS, BS-21, CAS, NRS, VDS; Version 5 = BS-21, FPS, VDS, NRS, CAS.

Data Analysis

Descriptive Statistics

Descriptive statistics were used for presenting the demographic characteristics of the subjects. The demographic characteristics were described in frequencies, percentages, means, and standard deviations. Face validity regarding preference, simplicity, and accuracy was also presented by descriptive statistics.

Inferential Statistics

Pearson product-moment correlation coefficient was used to assess concurrent validity, convergent validity, and test-retest reliability. However, Spearman rank correlation coefficient was also used especially when assessing the test-retest reliability of the VDS, the NRS, the FPS, and the BS-21 in the total subjects since the assumption of normality was violated.

Since the assumption of Chi-square that less than 20% of the cells should have the expected frequencies that are less than 5 was violated, Fisher's exact tests were used to test the differences of scale preference, simplicity, and accuracy of each pain scale among postoperative young adults, middle-aged adults, elderly without CI, and elderly with mild CI.

One-way ANOVA was used to test the differences of test-retest reliability coefficients of each pain scale among postoperative young adults,

middle-aged adults, elderly without CI, and elderly with mild CI. Kruskal-Wallis test was used to test the differences of concurrent validity coefficients and convergent validity coefficients of each pain scale among postoperative young adults, middle-aged adults, elderly without CI, and elderly with mild CI since the assumption of normality was violated in these situations.

For the interpretation of validity and reliability coefficients, the researcher used the following criteria.

Correlation Coefficients	Level of relationship	Level of validity/reliability
> .70	High	Very good
.40 - .69	Moderate	Good
.20 - .39	Low	Fair
< .20	Very low	Poor

CHAPTER 4

RESULTS AND DISCUSSIONS

This chapter presents two sections. The first section is the results of this study consisting of sample characteristics and psychometric properties of the pain intensity scales: the levels and the differences of (1) face validity, (2) concurrent validity, (3) convergent validity, and (4) test-retest reliability of each of the pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI. The second section is the discussion regarding the above results.

The Results of This Study

Sample Characteristics

Table 1 presents an overview of the demographic characteristics of the sample. This study included 200 adults with 50 in each group (young adults, middle-aged adults, elderly without CI, and elderly with mild CI). The mean age of patients was 55.56 years ($SD = 15.58$ years) with a range from 20 to 83 years. This sample comprised 108 male (54%) and 92 female (46%). For the educational level of the subjects, 3 patients (1.5%) did not have an education, 43 (21.5%) had a primary

school education, and 154 (77%) had a secondary school education or higher.

Table 1

Frequency and Percentage of Demographic Characteristics (N = 200)

Characteristics	n	%
Age (M = 55.56, SD = 15.58, Min = 20, Max = 83)		
20-44 yrs (young adults)	50	25
45-59 yrs (middle-aged adults)	50	25
≥ 60 yrs (elderly without CI)	50	25
≥ 60 yrs (elderly with mild CI)	50	25
Gender		
Male	108	54
Female	92	46
Religion		
Islam	9	4.5
Tibetan	1	0.5
No religion	190	95
Educational level		
No education	3	1.5
Elementary school	43	21.5
Secondary school	45	22.5
High school	47	23.5
Diploma or Bachelor degree	62	31
Occupation		
Government employee	22	11
Private employee	42	21
Unemployed	43	21.5
Retired	91	45.5
Student	2	1
Marital status		
Never married	8	4
Married	188	94
Divorced/separated	2	1
Widowed	2	1

Validity and Reliability of the Pain Intensity Scales

Face Validity

In this study, face validity of the pain intensity scales was assessed by three aspects including preference, simplicity, and accuracy. They are presented separately as follows.

For scale preference, nearly half of the subjects (42.5%) most preferred the FPS, followed by the VDS (29.5%) and the NRS (20%), whereas fewer subjects selected the CAS and the BS-21 as the most preferred scale (Table 2). The preference of the five pain scales in each group is presented in Table 3. The NRS ($n = 18$, 36%) was the most preferred scale in the young adult group. While, for the middle-aged adult, elderly without CI, and elderly with mild CI groups, the FPS was selected as the most preferred scale, 50% ($n = 25$), 40% ($n = 20$), and 48% ($n = 24$), respectively. Moreover, scale preference was significantly related to different age groups and the levels of CI (Fisher's exact: $p = .001$).

The pattern of scale simplicity was somewhat similar with the pattern of scale preference since most subjects who selected a specific scale as the most preferred also selected this scale as the simplest. In all subjects, the scales selected as simplest (Table 4) in order were as follows: the FPS (51.5%), the VDS (29%), the NRS (11%), the CAS (5.5%), and the BS-21 (3%). More specifically, in each of the four groups, the subjects selected the FPS as the simplest scale with a range from 50% to 52% (Table 5).

Furthermore, the four groups showed a significant difference in scale simplicity (Fisher's exact: $p = .006$).

For scale accuracy, the total number of scale inaccurate responses in all subjects was remarkably low (Table 6). The percentage of inaccurate responses for each scale in all subjects in order was as follows: the CAS (4.5%), the NRS (1.5%), the BS-21 (1%), the FPS (0.5%), and the VDS (0%). Specifically, the inaccurate responses for the CAS were due to its difficulty to understand for the participants, so they just left them blank; the inaccurate responses for the NRS, the BS-21, and the FPS included no rating on the scale, response falling between two numbers, and mistake in ordinal understanding of the scale. Interestingly, inaccurate responses were only found in middle-aged adult group and elderly with mild CI group. Moreover, Fisher's exact test was used to test whether each scale's accuracy was related to different age groups and the levels of CI (Table 7). Only the accuracy of the CAS was found to be significantly different among the four groups (Fisher's exact: $p = .007$) that elderly patients with mild CI group had the highest inaccurate responses.

Table 2
Preference of the Five Pain Intensity Scales in All Subjects (N = 200)

Scale	Level of preference				
	1 n (%)	2 n (%)	3 n (%)	4 n (%)	5 n (%)
VDS	59 (29.5)	73 (36.5)	58 (29)	8 (4)	2 (1)
NRS	40 (20)	31 (15.5)	44 (22)	74 (37)	11 (5.5)
FPS	85 (42.5)	49 (24.5)	31 (15.5)	17 (8.5)	18 (9)
CAS	9 (4.5)	42 (21)	39 (19.5)	32 (16)	78 (39)
BS-21	7 (3.5)	5 (2.5)	28 (14)	69 (34.5)	91 (45.5)

Table 3
Test of the Most Preferred Scales Classified by Different Age Groups and Levels of CI, Using Fisher's Exact Test (N = 200)

Scale	Selected as the most preferred				χ^2	p-value
	20-44 years n (%)	45-59 years n (%)	≥ 60 years without CI n (%)	≥ 60 years with mild CI n (%)		
VDS	9 (18)	16 (32)	12 (24)	22 (44)	29.79	.001
NRS	18 (36)	6 (12)	14 (28)	2 (4)		
FPS	16 (32)	25 (50)	20 (40)	24 (48)		
CAS	4 (8)	1 (2)	3 (6)	1 (2)		
BS-21	3 (6)	2 (4)	1 (2)	1 (2)		

Table 4
Simplicity of the Five Pain Intensity Scales in All Subjects (N = 200)

Scale	Level of simplicity				
	1 n (%)	2 n (%)	3 n (%)	4 n (%)	5 n (%)
VDS	58 (29)	79 (39.5)	53 (26.5)	8 (4)	2 (1)
NRS	22 (11)	35 (17.5)	54 (27)	78 (39)	11 (5.5)
FPS	103(51.5)	39 (19.5)	24 (12)	16 (8)	18 (9)
CAS	11 (5.5)	42 (21)	42 (21)	32 (16)	73 (36.5)
BS-21	6 (3)	5 (2.5)	27 (13.5)	66 (33)	96 (48)

Table 5
Test of the Simplest Scales Classified by Different Age Groups and Levels of CI, Using Fisher's Exact Test (N = 200)

Scale	Selected as the simplest				χ^2	p-value
	20-44 years n (%)	45-59 years n (%)	≥ 60 years without CI n (%)	≥ 60 years with mild CI n (%)		
	VDS	9 (18)	17 (34)	10 (20)		
NRS	9 (18)	4 (8)	9 (18)	-		
FPS	25 (50)	26 (52)	26 (52)	26 (52)		
CAS	5 (10)	1 (2)	4 (8)	1 (2)		
BS-21	2 (4)	2 (4)	1 (2)	1 (2)		

Table 6
Accuracy of Five Pain Intensity Scales in All Subjects (N = 200)

Scale	Accurate response n (%)	Inaccurate response n (%)
VDS	200 (100)	0 (0)
NRS	197 (98.5)	3 (1.5)
FPS	199 (99.5)	1 (0.5)
CAS	191 (95.5)	9 (4.5)
BS-21	198 (99)	2 (1)

Table 7
Test of Scale Accuracy Classified by Different Age Groups and Levels of CI, Using Fisher's Exact Test (N = 200)

Scale	Accuracy				χ^2	p-value
	20-44 years n (%)	45-59 years n (%)	≥ 60 years without CI n (%)	≥ 60 years with mild CI n (%)		
VDS accurate	50 (100)	50 (100)	50 (100)	50 (100)		
Inaccurate	-	-	-	-		
NRS accurate	50 (100)	49 (98)	50 (100)	48 (96)	2.91	.62
inaccurate	-	1 (2)	-	2 (4)		
FPS accurate	50 (100)	49 (98)	50 (100)	50 (100)	2.82	1.00
inaccurate	-	1 (2)	-	-		
CAS accurate	50 (100)	47 (94)	50 (100)	44 (88)	10.05	.007
inaccurate	-	3 (6)	-	6 (12)		
BS-21 accurate	50 (100)	49 (98)	50 (100)	49 (98)	2.13	1.00
inaccurate	-	1 (2)	-	1 (2)		

Concurrent Validity

For concurrent validity of the five pain scales (Table 8), Pearson

product-moment correlation coefficients between the scores of the modified Pain Interference Scale (PIS) and the scores of each pain scale ranged from .72 to .77 in all subjects, indicating very good concurrent validity. Specifically, the concurrent validity of each scale across the four groups was as follows: the VDS ($r = .72 - .85$), the NRS ($r = .58 - .83$), the FPS ($r = .62 - .81$), the CAS ($r = .65 - .82$), and the BS-21 ($r = .68 - .81$). The concurrent validity of all the five pain scales in each group was also presented with Pearson product-moment correlation coefficients ranging from .72 to .80 in young adult group; .73 to .85 in the middle-aged adult group; .72 to .82 in elderly without CI group; and a slightly reduction of .58 to .73 in elderly with mild CI group. In addition, the Kruskal-Wallis tests indicated that the concurrent validity coefficients of each pain scale were not significantly different among the four groups (Table 9).

Table 8

Concurrent Validity: Pearson Product-moment Correlation Coefficients Between the Scores of the Modified Pain Interference Scale and Scores of Each of the Five Pain Intensity Scales in All Subjects (listwise deletion N = 188)

Scale	Concurrent validity (correlation coefficient: r)	p-value	Level of concurrent validity
VDS	.77	.000	Very good
NRS	.75	.000	Very good
FPS	.72	.000	Very good
CAS	.74	.000	Very good
BS-21	.73	.000	Very good

Note. The deleted cases were due to that the patients had errors in using the pain scales.

Table 9

Test of the Difference of Concurrent Validity Coefficients Comparing Among Different Age Groups and Levels of CI, Using Kruskal-Wallis Test (listwise deletion N = 188)

Scale	Level of concurrent validity correlation coefficient (r)				χ^2	p-value
	20-44 years (n = 50)	45-59 years (n = 46)	≥ 60 years without CI (n = 50)	≥ 60 years with mild CI (n = 42)		
VDS	.72	.85	.78	.73	1.26	.74
NRS	.80	.83	.80	.58	1.06	.79
FPS	.74	.81	.72	.62	.48	.92
CAS	.73	.73	.82	.65	.33	.95
BS-21	.76	.81	.80	.68	.74	.86

Note. The deleted cases were due to that the patients had errors in using the pain scales.

Convergent Validity

Convergent validity of the four scales (NRS, FPS, CAS, and BS-21) was proved to be very good since Pearson product-moment correlation coefficients between the scores of the VDS and the scores of each of the four pain scales ranged from .84 to .90 in all subjects (Table 10). Since the NRS is another tool that is easily to understand and is recommended for pain assessment in elderly with cognitive function ranging from none to mild or moderate CI from a study of literature review (Hadjistavropoulos et al., 2007), therefore, if the correlation between the NRS and the VDS is high enough ($> .70$), it will provide evidence for convergent validity of both the NRS and the VDS. Therefore, the Pearson product-moment correlation coefficient .90 between the NRS and the VDS in all subjects supported the convergent validity of both of them. Specifically, the results showed that the convergent validity of each of the four pain scales across the four groups was very good with Pearson product-moment correlation coefficients ranging from .83 to .95 for the NRS, .81 to .92 for the FPS, .85 to .89 for the CAS, and .81 to .95 for the BS-21 (Table 11). In addition, the convergent validity of all the pain scales in each group was proved to be very good. Furthermore, A Kruskal-Wallis test indicated that the convergent validity coefficients of the NRS were not significantly different among the four groups ($\chi^2 = 2.07, p = .56$). This was also the case for the convergent validity coefficients of the FPS ($\chi^2 = 1.55, p = .67$), the CAS ($\chi^2 = 0.32, p = .96$), and the BS-21 ($\chi^2 = 2.08, p = .56$).

Table 10

Convergent Validity: the Pearson Product-moment Correlation Coefficients Between the Scores of the VDS and the Scores of Each of the Four Pain Intensity Scales in All Subjects (listwise deletion $N = 188$)

Scale	Convergent validity correlation coefficient (r)	p-value	Level of convergent validity
NRS	.90	.000	Very good
FPS	.84	.000	Very good
CAS	.86	.000	Very good
BS-21	.89	.000	Very good

Note. The deleted cases were due to that the patients had errors in using the pain scales.

Table 11

Test of the Difference of Convergent Validity Coefficients Comparing Among Different Age Groups and Levels of CI, Using Kruskal-Wallis Test (listwise deletion $N = 188$)

Scale	Level of convergent validity correlation coefficient (r)				χ^2	p-value
	20-44 years (n = 50)	45-59 years (n = 46)	≥ 60 years without CI (n = 50)	≥ 60 years with mild CI (n = 42)		
NRS	.90	.95	.92	.83	2.07	.56
FPS	.83	.92	.81	.81	1.55	.67
CAS	.86	.88	.89	.85	0.32	.96
BS-21	.89	.95	.93	.81	2.08	.56

Note. The deleted cases were due to that the patients had errors in using the pain scales.

Test-retest Reliability

The test-retest reliability of the five scales, assessed by Spearman rank correlation coefficients (for the VDS, the NRS, the FPS, and the BS-21) and Pearson product-moment correlation coefficient (for the CAS) between the 3-day vividly

remembered pain ratings, ranged from .75 to .80 in all subjects (Table 12). Specifically, the test-retest reliability of each pain scale across the four groups with Pearson product-moment correlation coefficients ranged from .64 to .93 for the VDS, .62 to .89 for the NRS, .63 to .87 for the FPS, .65 to .85 for the CAS, and .59 to .88 for the BS-21. In addition, the test-retest reliability of all the pain scales ranged from .85 to .93 in young adult group, .75 to .83 in middle-aged adult group, .65 to .87 in elderly without CI group, and .59 to .65 in elderly with mild CI group where there was a slight decrease. Furthermore, the test-retest reliability coefficients of each pain scale were not significantly different by the four groups (Table 13).

Table 12

Test-retest Reliability: Correlation Coefficients Between the Vividly Remembered Pain Ratings at a 3-day Interval in All Subjects (listwise deletion N = 153)

Scale	Test-retest reliability correlation coefficient	p-value	Level of test-retest reliability
VDS	.80 ^a	.000	Very good
NRS	.79 ^a	.000	Very good
FPS	.76 ^a	.000	Very good
CAS	.75 ^b	.000	Very good
BS-21	.76 ^a	.000	Very good

Note. The deleted cases were due to that the patients had errors in using the pain scales or they did not have a recalled pain that could be used for testing the reliability; ^aSpearman rank correlation coefficient; ^bPearson product-moment correlation coefficient.

Table 13

Test of the Difference of Test-retest Reliability Coefficients Comparing Among Different Age Groups and Levels of CI, Using one-way ANOVA (listwise deletion N = 153)

Scale	Level of test-retest reliability correlation coefficient (r)				F	p-value
	20-44 years (n = 43)	45-59 years (n = 39)	≥ 60 years without CI (n = 40)	≥ 60 years with mild CI (n = 31)		
VDS	.93	.75	.87	.64	0.35	.79
NRS	.89	.80	.83	.62	0.52	.67
FPS	.87	.81	.73	.63	1.30	.28
CAS	.85	.83	.65	.65	2.43	.07
BS-21	.88	.81	.80	.59	1.23	.30

Note. The deleted cases were due to that the patients had errors in using the pain scales or they did not have a recalled pain that could be used for testing the reliability.

Discussion of the Study Results

Findings from this study demonstrate that all five pain scales are reliable and valid for assessing pain in Chinese adults including elderly with mild CI. These findings are consistent with the study by Herr et al. (2004), who reported that the young and old adults including those with mild to moderate CI were able to use the selected pain intensity scales. In addition, the present findings also confirm the previous studies reporting that the elderly with none to moderate CI were generally able to use the pain intensity scales (Chibnall & Tait, 2001; Closs et al., 2004; Samabub, Petpichetchian, & Kitrungrrote, 2009; Taylor & Herr, 2003; Ware et al.,

2006).

Face Validity of the Pain Intensity Scales

Face validity should be considered for selecting a right scale especially when the scales are all good at other aspects of validity and reliability. The results of present study revealed that the FPS was ranked best in face validity across the four groups as nearly half of the patients selected it as both the most preferred and simplest scale and it had low inaccurate responses. The reason might be that completing the FPS does not require reading, writing, and expressive ability, making it applicable for the adults, especially the older adults who had difficulty to communicate their pain by traditional scales (Herr et al., 1998). The face validity of the NRS and the VDS was similar and ranked following the FPS since they are simple and easy to understand. However, the BS-21 and the CAS were ranked last in face validity. Few subjects selected them as the most preferred and simplest scale. In addition, the CAS had the highest inaccurate responses and the reason might be that completing the CAS needs more abstract thinking ability compared with other scales. This is consistent with the study by Closs et al. (2004) who suggested that the CAS appeared to be conceptually more difficult to understand compared with the VDS, the FPS, and the NRS. In contrast, the findings from Scherder and Bouma (2000) showed that the CAS worked best and was correctly interpreted by all the elderly without dementia and with early AD, compared with the FPS and the Facial Affective Scale (FAS) which is much suitable for children because of the childlike facial depictions. In terms of face validity,

overall, the findings from the present study are similar to the study by Li et al. (2007), who found that 48.1% of the Chinese adults preferred the FPS-R and the FPS-R had low error rates followed by the VDS and the NRS. The results of the present study are also similar to the previous study (Taylor & Herr, 2003) which compared the VDS, the NRS, the FPS, and the Iowa Pain Questionnaire (IPT) and reported that the FPS was the most preferred scale and had no failures in African-American elderly including those with CI. In contrast, another two studies compared the similar scales in a primarily Caucasian sample, they found that the NRS and the VDS were the most preferred scales and had low errors in young and old adults without and with CI (Herr et al., 2004; Taylor et al., 2005), suggesting that the ethnic or culture differences might have an effect on scale preference.

The findings regarding the age and CI differences on scale preference and scale accuracy are generally inconsistent between the present study and the previous studies. For scale preference, the present study found that scale preference was significantly related to different age groups and the levels of CI. Consistently, Peters et al. (2007) found an age difference on scale preference with patients of 75 years or older more preferred the VDS ($p = .014$). However, some other studies found that scale preference was not related to age (Gagliese et al., 2005; Herr et al., 2004; Li et al., 2007) and cognitive status (Herr et al.). For scale accuracy, the present study found that the accuracy of the CAS was significantly different across age groups and the levels of CI. Unlike the parallel study conducted in Thai population by Samabub et

al. (2009), they found no significant difference across age groups and the levels of CI in using the CAS. There were two differences related to the CAS itself between the present study and the Samabub et al.'s study: the alignment and the color. In the present study, the CAS was aligned vertically and the light yellow hue was used at the bottom "no pain" to the deep red hue at the top "worst pain". For Samabub et al.'s study, the alignment was horizontal and the colors were ranged from yellow-green "no pain" to dark brown "worst pain". These differences may contribute to the different findings. However, some studies also reported that age was significantly related to scale inaccurate responses. Gagliese et al. found that the patients who made errors on the VAS-horizontal (VAS-H) were older than those who did not make errors ($p \leq .009$) and Peters et al. reported that age was proved to be significantly related to making a mistake, with older patients making more mistakes ($p = .02$) on the pain scales. In contrast, Herr et al. found that age did not impact failure to use the pain scale but cognitive and psychomotor impairment did since the cognitive and motor impairment were found to significantly increase the relative risk of failure to use the VAS successfully ($p < .05$).

Concurrent Validity of the Pain Intensity Scales

The concurrent validity of all five scales in each group was supported and ranged from good to very good ($r = .58 - .85$) across the four groups. In addition, the findings from the present study also provided evidence that the concurrent validity

coefficients of each scale were not significantly different among different age groups and the levels of CI, indicating that all the five scales were equally valid regarding concurrent validity. Similarly, the study with Thai population found no significant difference of the five pain scales with respect to concurrent validity coefficients among different age groups and the levels of CI (Samabub e al., 2009). However, in that study, the concurrent validity coefficients ($r = .07 - .64$) were lower than what were found in this present study in which the researchers found the lower coefficients in elderly with mild CI.

Convergent Validity of the Pain Intensity Scales

Findings of this study demonstrated that the convergent validity of all the five pain scales in each group was strongly supported and ranged from .81 to .95 across the four groups. It is also important to note that the results from the present study showing that the convergent validity coefficients of each scale were not significantly different across the four groups, indicating that all the pain scales were similarly valid across the subjects with respect to the convergent validity. These are consistent with the following six studies. The first study by Samabub et al. (2009) was almost identical to this study. They found that among Thai postoperative patients, the convergent validity coefficients across the four groups ranged from .25 to .71 and the lowest coefficient was found in the correlation between the VDS and the BS-21 in the elderly with mild CI group ($r = .25$). The second study by Li et al. (2007) examined the

pain scales including the VAS, the NRS, the VDS, and the FPS-R in Chinese postoperative adults and found that the correlations between the four scales for rating the worst pain were strong with Spearman rank correlation coefficients ranging from .80 to .99. The third and fourth studies (Taylor et al., 2005; Taylor & Herr, 2003) found the good to very good inter-tool correlations ($\rho = .48 - .97$) among the NRS, the VDS, the FPS, and the IPT in elderly without and with CI. Similarly, the fifth study by Ware et al. (2006) supported the inter-tool correlations ($\rho = .56 - .90$) of the NRS, the VDS, the FPS-R, and the IPT in using with elderly without and with CI. Consistently, the sixth study conducted by Gagliese et al. (2005) reported good to very good convergent validity of the NRS, the VDS, and the VAS in young adults ($r = .60 - .93$) and old adults ($r = .72 - .91$) and the convergent validity of the pain scales did not differ between age groups. However, the present findings are also contrary to the previous study (Gagliese & Katz, 2003) reporting that age differences in the convergent validity were evident since the correlation between the VAS and the McGill Pain Questionnaire (MPQ) scores was significantly lower in the old adult group than young adult group. This discrepancy might be explained by the methodological differences in evaluating convergent validity between Gagliese and Katz's study and the present study. The present study used the unidimensional tool VDS to correlate with other pain scales to assess convergent validity similarly to Samabub et al.'s study, whereas, Gagliese and Katz correlated the pain scales with the multidimensional tool, the MPQ, which might have inadvertently introduced fatigue effects especially for the

older group. Importantly, when Closs et al. (2004) examined the NRS, the VDS, the FPS, the CAS, and the mechanical visual analogue scale (MVAS) for use in patients with different levels of CI, they reported that the Spearman rank correlation coefficients between scores on the pain scales were acceptable for elderly without CI (.50 - .68), with mild CI (.62 - .77), and with moderate CI (.38 - .88) but poor for elderly with severe CI (-.09 - .68) which was due to that the elderly with severe CI could not use the pain scales meaningfully.

Test-retest Reliability of the Pain Intensity Scales

Reliability is also an important criterion when selecting a scale. In the present study, the test-retest reliability of all the five pain scales in each group was supported and ranged from good to very good ($r = .59 - .93$) across the four groups with a slight reduction in the elderly with mild CI which might be due to the memory impairment in this group as noted by others. In addition, the findings from the present study suggested that there was no significant difference in the test-retest reliability coefficients of each scale among different age groups and the levels of CI. Repeatedly, this finding is consistent with the study by Samabub et al. (2009) which found that the reliability coefficients ranged from .59 to .86 across the four groups. In these two studies, a 3-day interval rather than the commonly used 2-week interval was used. To our knowledge, it is reasonable to select a 3-day interval since the test-retest reliability over a 2-week interval was found to be unacceptable ($\rho = .26 - .67$) when Taylor et al.

(2005) examined the similar scales (VDS, NRS, FPS, and IPT) in the elderly with CI group and Chibnall and Tait (2001) suggested a memory degrade at approximately 3 days in patients with CI. However, in some studies, the test-retest reliability at a 2-week interval was also acceptable with Spearman rank correlation coefficients ranging from .52 to .83 for the VDS, the NRS, the FPS, and the IPT (Taylor & Herr, 2003) and .76 to .89 for the NRS, the VDS, the FPS-R, and the IPT (Ware et al., 2006) in both elderly without and with CI. In addition, in an earlier study for evaluating the FPS in use with the elderly without CI, the researchers found that the test-retest reliability of the FPS over a 2-week interval was strong ($\rho = .94$) (Herr et al., 1998). Therefore, with the similar findings from several studies including the present study, it can be concluded that using a 3-day to 2-week interval for testing test-retest reliability in the elderly with mild CI can be accepted.

It is also notable that, on the 2nd postoperative day, 97% of the patients experiencing pain in the past 24 hours and the mean score of average pain in the past 24 hours was 4 at a moderate level according to the NRS. This further proves that the postoperative pain is still undertreated in Chinese patients.

In summary, the results from the present study demonstrated that the face validity, concurrent validity, convergent validity, and test-retest reliability of the five selected scales were strongly supported in use with the four groups including young adults, middle-aged adults, elderly without CI, and elderly with mild CI. In addition, there was little evidence for validity and reliability differences on the four

groups, suggesting that these five scales were standard across the four groups. Specifically, the only differences in psychometric indices among the four groups were found in face validity. Therefore, although all the five pain scales were psychometrically sound in use with the four groups, face validity suggested that the FPS was the first choice for assessing pain intensity across the four groups as it was selected as the most preferred and simplest scale as well as its high accuracy compared to other scales. The VDS and the NRS would be the second choice and the BS-21 and the CAS were the last choice since they were less preferred by the patients and had relatively high inaccurate responses especially for the CAS.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

This chapter presents the summary of the study findings, strengths and limitations of this study, and implications and recommendations for future studies.

Summary of the Study Findings

This study was aimed to examine the psychometric properties of five evidence-supported pain intensity scales, including the VDS, the NRS, the FPS, the CAS, and the BS-21, among postoperative young adult patients, middle-aged adult patients, elderly patients without CI, and elderly patients with mild CI. In terms of face validity, the FPS was ranked best across the four groups as nearly half of the patients selected it as both the most preferred and simplest scale and it had low inaccurate responses. The face validity of the VDS and the NRS was similar and ranked following the FPS since they are simple and easy to understand. However, the BS-21 and the CAS were ranked last in face validity since few patients selected them as the most preferred and simplest scale. In addition, scale simplicity and preference were significantly related to different age groups and the levels of CI. Moreover, the accuracy of the CAS was significantly related to different age groups and the levels of CI that elderly patients with mild CI group had the highest inaccurate responses. Concurrent validity, convergent validity, and test-retest reliability of the five selected

scales were strongly supported in use with the four groups. In addition, the findings from the present study also showed that the concurrent validity, convergent validity, and test-retest reliability of each scale were not significantly different among different age groups and the levels of CI, indicating all the five scales were equally valid regarding these three aspects of validity. Overall, the findings of this study support the validity and reliability of all five scales (VDS, NRS, CAS, FPS, and BS-21) for pain assessment in Chinese adults including elderly with mild CI. However, the FPS is proposed as the best scale for pain assessment in Chinese patients because of its good validity and reliability as well as highest face validity followed by the VDS and the NRS.

Strengths and Limitations of the Study

There are two main strengths in this study. For one thing, we extended the evidence-supported pain intensity scales to a broader spectrum of patients including the elderly with mild CI compared to the previous Chinese study by Li et al. (2007). For another, we randomized the scale administration order to control order effect by making five different versions of the questionnaire. However, we should be careful in generalizing the findings since the present study has several limitations that should be mentioned. Firstly, purposive sampling technique used in this study could limit representativeness of the population. However, the researchers tried to recruit all those patients who met the inclusion criteria during the data collection period. We also

ensured that each participant could randomly receive one of the five different versions of the questionnaire that would counterbalance the order effect. Secondly, although this study extended the selected pain intensity scale to the heterogeneous groups including elderly with mild CI, the patients with moderate CI were not included. Since some studies supported that patients with mild to moderate CI could rate pain reliably and validity (Chibnall & Tait, 2001; Closs et al., 2004), future studies are needed to examine whether these pain scales can be generalized to the moderate CI Chinese patients undergoing surgery. Thirdly, although this study investigated the preliminary psychometric properties of the pain intensity scales, scale sensitivity to detect changes in pain sensation which is also an important criterion in both clinical practice and research has yet to be determined and needs to be explored in future studies. Finally, this study was conducted in Kunming which might limit the generalizability of these findings.

Implications and Recommendations

Implications

The findings in this study provide important implications and insights for nursing education, practice, and research.

Nursing Education

The nurses should be aware and recognize pain assessment as a prerequisite for effective pain management. The evidence for pain assessment in this study can be added in the nursing teaching curriculum and program. Therefore, the nursing educators can transfer this evidence to the students or nurses to update their knowledge in postoperative pain assessment across different age groups and the levels of CI.

Nursing Practice

The findings of this study support the validity and reliability of all five scales (VDS, NRS, CAS, FPS, and BS-21) for pain assessment in Chinese adults including elderly with mild CI. In addition, the FPS is proposed as the best scale followed by the VDS and the NRS. Therefore, the nurses can use these findings to select the best pain assessment tool for Chinese patients. Consequently, the patients can get effective pain management. Moreover, considering the common visual impairment in elderly, the present study used enlarged print and darkened lines to enhance visualization; therefore, it is important to incorporate these adaptations into tool development (Herr et al., 2004) as well as facilitate adequate lighting and hearing devices (Taylor & Herr, 2003) when assessing pain in elderly.

Nursing Research

The present findings add new knowledge to a body of accumulated data reference on research related to evaluating psychometric properties of pain intensity scales in postoperative patients. It also provides a knowledge research base for nurses or other health care professionals for future research development or replicating similar study in their culture.

Recommendations for Future Research

Firstly, considering the limitations in this study, future studies are needed to determine whether the results can be generalized to the elderly with moderate CI in Chinese population. In addition, it is also important to identify the appropriate pain assessment tools for other population such as chronic pain and cancer pain patients. Secondly, other aspects of psychometric properties especially the scale sensitivity needs to be explored in future studies. Thirdly, since the Chinese patients assessed pain more accurately with the vertical version of the VAS than with the commonly used horizontal version, further study for testing all scales in vertical alignment may be needed. Fourthly, the gender and educational level differences on psychometric properties of pain scales may need to be explored in further study. Finally, the culture issues that might contribute to the different findings need to be explored. In addition, considering the culture differences, replication study in other culture is also recommended.

REFERENCES

- Ahles, T. A., Blanchard, E. B., & Ruckdeschel, J. C. (1983). The multidimensional nature of cancer-related pain. *Pain, 17*, 277-288.
- Albert, M. S., & Blacker, D. (2006). Mild cognitive impairment and dementia. *Annual Review of Clinical Psychology, 2*, 379-388.
- Apfelbaum, J. L., Chen, C., Mehta, S. S., & Gan, T. J. (2003). Postoperative pain experience: Results from a national survey suggest postoperative pain continues to be undermanaged. *Anesthesia and Analgesia, 97*, 534-540.
- Aronow, W. S. (2003). Effects of aging on the heart: Cardiovascular disease. In R. C. Tallis & H. M. Fillit (Eds.), *Brocklehurst's textbook of geriatric medicine and gerontology* (6th ed.) (pp. 341-348). London: Churchill Livingstone.
- Aubrun, F., & Marmion, F. (2007). The elderly patient and postoperative pain treatment. *Best Practice & Research, 21*, 109-127.
- Aun, C., Lam, Y. M., & Collett, B. (1986). Evaluation of the use of visual analogue scale in Chinese patients. *Pain, 25*, 215-221.
- Berkley, K. J., & Holdcroft, A. (1999). Sex and gender differences in pain. In P. Wall & R. Melzack (Eds.), *Textbook of pain* (4th ed.) (pp. 951-965). London: Churchill Livingstone.
- Bieri, D., Reeve, R. A., Champion, G. D., Addicoat, L., & Ziegler, J. B. (1990). The Faces Pain Scale for the self-assessment of the severity of pain experienced by

- children: Development, initial validation, and preliminary investigation for ratio scale properties. *Pain*, *41*, 139-150.
- Breivik, H. (2002). How to implement an acute pain service. *Best Practice & Research*, *16*, 527-547.
- Brunner, L. S., & Suddarth, D. S. (2000). *Textbook of medical-surgical nursing* (9th ed.). Philadelphia: J.B. Lippincott.
- Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. In J. L. Sullivan & R. G. Niemi (Eds.), *Series: Quantitative applications in the social sciences* (pp. 9-51). Beverly Hills, CA: Sage Publications.
- Chakour, M. C., Gibson, S. J., Bradbeer, M., & Helme, R. D. (1996). The effect of age on A delta- and C-fibre thermal pain perception. *Pain*, *64*, 143-152.
- Chen, L. M., Miaskowski, C., Dodd, M., & Pantilat, S. (2008). Concepts within the Chinese culture that influence the cancer pain experience. *Cancer Nursing*, *31*, 103-108.
- Chen, P. P., Chui, P. T., Ma, M., & Gin, T. (2001). A prospective survey of patients after cessation of patient-controlled analgesia. *Anesthesia and Analgesia*, *92*, 224-227.
- Chibnall, J. T., & Tait, R. C. (2001). Pain assessment in cognitively impaired and unimpaired older adults: a comparison of four scales. *Pain*, *92*, 173-186.
- Chung, J. W., & Lui, J. C. (2003). Postoperative pain management: Study of patients' level of pain and satisfaction with health care providers' responsiveness to their reports of pain. *Nursing & Health Sciences*, *5*, 13-21.

- Cleeland, C. S. (1984). The impact of pain on the patient with cancer. *Cancer, 54*, 2635-2641.
- Closs, S. J., Mphil, Barr, B., Briggs, M., Cash, K., & Seers, K. (2004). A comparison of five pain assessment scales for nursing home residents with varying degrees of cognitive impairment. *Journal of Pain and Symptom Management, 27*, 196-205.
- Commenges, D., Gagnon, M., Letenneur, L., Dartigues, J. F., Barberger-Gateau, P., & Salamon, R. (1992). Improving screening for dementia in the elderly using, Mini-Mental State Examination subscores, Benton's Visual Retention Test, and Isaacs' Set Test. *Epidemiology, 3*, 185-188.
- Craig, D. B. (1981). Postoperative recovery of pulmonary function. *Anesthesia and Analgesia, 60*, 46-52.
- Dahl, J. L. (1996). Effective pain management in terminal care. *Clinics in Geriatric Medicine, 12*, 279-300.
- Dawson, P. (1998). Cognitively impaired residents receive less pain medication than non-cognitively impaired residents. *Perspectives, 22*(4), 16-17.
- Defrin, R., Lotan, M., & Pick, C. G. (2006). The evaluation of acute pain in individuals with cognitive impairment: A differential effect of the level of impairment. *Pain, 124*, 312-320.
- Douglas, M. (1999). Pain as the fifth vital sign: will cultural variations be considered? *Journal of Transcultural Nursing 10*, 285.
- Dunwoody, C. J., Krenzischek, D. A., Pasero, C., Rathmell, J. P., & Polomano, R. C.

- (2008). Assessment, physiological monitoring, and consequences of inadequately treated acute pain. *Journal of Perianesthesia Nursing*, 23, 15-27.
- Duppils, G. S., & Wikblad, K. (2000). Acute confusional states in patients undergoing hip surgery: A prospective observation study. *Gerontology*, 46(1), 36-43.
- Feldt, K. S., Warne, M. A., & Ryden, M. B. (1998). Examining pain in aggressive cognitively impaired older adults. *Journal of Gerontological Nursing*, 24(11), 14-22.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Gagliese, L., & Katz, J. (2003). Age differences in postoperative pain are scale dependent: A comparison of measures of pain intensity and quality in younger and older surgical patients. *Pain*, 103, 11-20.
- Gagliese, L., & Melzack, R. (2006). Pain in elderly. In S. B. McMahon & M. Koltzenburg (Eds.), *Textbook of pain* (5th ed.) (pp. 1169-1177). London: Churchill Livingstone.
- Gagliese, L., Weizblit, N., Ellis, W., & Chan, V. W. (2005). The measurement of postoperative pain: A comparison of intensity scales in younger and older surgical patients. *Pain*, 117, 412-420.
- Hadjistavropoulos, T., & Craig, K. D. (2002). A theoretical framework for understanding self-report and observational measures of pain: A communications

- model. *Behaviour Research and Therapy*, 40, 551-570.
- Hadjistavropoulos, T., Herr, K., Turk, D. C., Fine, P. G., Dworkin, R. H., Helme, R., et al. (2007). An interdisciplinary expert consensus statement on assessment of pain in older persons. *The Clinical Journal of Pain*, 23(1 Suppl), S1-43.
- Heffline, M. S. (1990). Exploring nursing interventions for acute pain in the postanesthesia care unit. *Journal of Post Anesthesia Nursing*, 5, 321-328.
- Herr, K. A., Mobily, P. R., Kohout, F. J., & Wagenaar, D. (1998). Evaluation of the Faces Pain Scale for use with the elderly. *The Clinical Journal of Pain*, 14, 29-38.
- Herr, K. A., Spratt, K., Mobily, P. R., & Richardson, G. (2004). Pain intensity assessment in older adults: use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. *The Clinical Journal of Pain*, 20, 207-219.
- Hicks, C. L., von Baeyer, C. L., Spafford, P. A., van Korlaar, I., & Goodenough, B. (2001). The Faces Pain Scale-Revised: Toward a common metric in pediatric pain measurement. *Pain*, 93, 173-183.
- Ho, K., Spence, J., & Murphy, M. F. (1996). Review of pain-measurement tools. *Annals of Emergency Medicine*, 27, 427-432.
- Hurley, A. C., Volicer, B. J., Hanrahan, P. A., Houde, S., & Volicer, L. (1992). Assessment of discomfort in advanced Alzheimer patients. *Research in Nursing & Health*, 15, 369-377.
- International Association for the Study of Pain (IASP). (1994). Retrieved January 27,

- 2009, from [http:// www. Iasp-pain. Org/terms-p.html](http://www.iasp-pain.Org/terms-p.html).
- Jacobson, S. F. (2004). Measuring pain. In M. F. Stromborg & S. J. Olsen (Eds.), *Instruments for clinical health-care research* (3rd ed.) (pp. 603-644). Boston: Jones and Bartlett.
- Jensen, M. P., Miller, L., & Fisher, L. D. (1998). Assessment of pain during medical procedures: A comparison of three scales. *The Clinical Journal of Pain, 14*, 343-349.
- Karanikolas, M., & Swarm, R. A. (2000). Current trends in perioperative pain management. *Anesthesiology Clinics of North America, 18*, 575-599.
- Keogh, E., & Herdenfeldt, M. (2002). Gender, coping and the perception of pain. *Pain, 97*, 195-201.
- Knapp, T. R. (1998). *Quantitative nursing research*. London: Thousand Oaks: SAGE.
- Kovach, C. R., Weissman, D. E., Griffie, J., Matson, S., & Muchka, S. (1999). Assessment and treatment of discomfort for people with late-stage dementia. *Journal of Pain and Symptom Management, 18*, 412-419.
- Kukull, W. A., Larson, E. B., Teri, L., Bowen, J., McCormick, W., & Pfanschmidt, M. L. (1994). The Mini-Mental State Examination score and the clinical diagnosis of dementia. *Journal of Clinical Epidemiology, 47*, 1061-1067.
- Lasch, K. E. (2000). Culture, pain, and culturally sensitive pain care. *Pain Management Nursing, 1*, 16-22.
- Li, L., Liu, X., & Herr, K. (2007). Postoperative pain intensity assessment: A

comparison of four scales in Chinese adults. *Pain Medicine*, 8, 223-234.

Li, J. P., Bai, Q., Liu, S. Z. & Li, X. M. (2000). *Community nursing*. Beijing: People's Health Press (in Chinese).

Li, M., Ng, T. P., Kua, E. H., & Ko, S. M. (2006). Brief informant screening test for mild cognitive impairment and early Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 21, 392-402.

Loan, W. B., & Morrison, J. D. (1967). The incidence and severity of postoperative pain. *British Journal of Anaesthesia*, 39, 695-698.

Lou, M. F., Dai, Y. T., Huang, G. S., & Yu, P. J. (2003). Postoperative cognitive changes among older Taiwanese patients. *Journal of Clinical Nursing*, 12, 579-588.

Lou, M. F., Dai, Y. T., Huang, G. S., & Yu, P. J. (2007). Identifying the most efficient items from the Mini-Mental State Examination for cognitive function assessment in older Taiwanese patients. *Journal of Clinical Nursing*, 16, 502-508.

Lynch, E. P., Lazor, M. A., Gellis, J. E., Orav, J., Goldman, L., & Marcantonio, E. R. (1997). Patient experience of pain after elective noncardiac surgery. *Anesthesia and Analgesia*, 85, 117-123.

Macintyre, P. E., & Ready, L. B. (2002). *Acute pain management* (2nd ed.). WB Saunders: Edinburgh.

Mackintosh, C. (2007). Assessment and management of patients with post-operative pain. *Nursing Standard*, 22, 49-55.

- Mariani, E., Monastero, R., & Mecocci, P. (2007). Mild cognitive impairment: A systematic review. *Journal of Alzheimer's disease, 12*, 23-35.
- Martinelli, A. M. (1987). Pain and ethnicity. How people of different cultures experience pain. *AORN Journal, 46*, 273-274, 276, 278 passim.
- McCaffery, M. (1981). *Nursing mangement of the patient with pain* (3rd ed.). Philadelphia: J.B. Lippincott.
- McCaffery, M., & Beebe, A. (1994). *Pain clinical manual for nursing practice*. London: Mosby.
- McCaffery, M., & Pasero, C. (1999). *Pain clinical manual* (2nd ed.). MO: Mosby.
- Mc Donald, M. E. (2002). *Systemic assessment of learning outcomes: Developing multiple-choice exams*. Boston: Jones and Bartlett.
- McGrath, P. A., Seifert, C. E., Speechley, K. N., Booth, J. C., Stitt, L., & Gibson, M. C. (1996). A new analogue scale for assessing children's pain: An initial validation study. *Pain, 64*, 435-443.
- McGrath, P. J., Rosmus, C., Canfield, C., Campbell, M. A., & Hennigar, A. (1998). Behaviours caregivers use to determine pain in non-verbal, cognitively impaired individuals. *Developmental Medicine and Child Neurology, 40*, 340-343.
- McGuire, D. B. (1992). Comprehensive and multidimensional assessment and measurement of pain. *Journal of Pain and Symptom Management, 7*, 312-319.
- McGuire, D. B., Kim, H. J., & Lang, X. T. (2004). Measuring pain. In M. F. Stromborg & S. J. Olsen (Eds.), *Instruments for clinical health-care research* (3rd ed.) (pp.

- 603-644). Boston: Jones and Bartlett.
- Morris, J. C., & Cummings, J. (2005). Mild cognitive impairment (MCI) represents early-stage Alzheimer's disease. *Journal of Alzheimer's Disease*, 7, 235-239; discussion 255-262.
- Munro, B. H. (2001). *Statistical methods for health care research*. In B. H. Munro (Eds.), *Correlation* (4th ed.) (pp. 223-244). Philadelphia: Lippincott.
- Nendick, M. (2000). Patient satisfaction with post-operative analgesia. *Nursing Standard*, 14, 32-37.
- Nunnally & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Okamoto, Y., Case, R., Bleiker, C., & Henderson, B. (1996). Cross-cultural investigations. *Monographs of the Society for Research in Child Development*, 61, 131-155.
- Pasero, C. (2003). Pain in the critically ill patient. *Journal of Perianesthesia Nursing*, 18, 422-425.
- Pasero, C. L., & McCaffery, M. (1996). Managing postoperative pain in the elderly. *The American Journal of Nursing*, 96, 38-45.
- Peeters-Asdourian, C. (2002). *Acute pain mangement* (2nd ed.). Philadelphia: Lippincott Williams & Wilkins.
- Pesonen, A., Suojaranta-Ylinen, R., Tarkkila, P., & Rosenberg, P. H. (2008). Applicability of tools to assess pain in elderly patients after cardiac surgery. *Acta*

Anaesthesiologica Scandinavica, 52, 267-273.

Peters, M. L., Patijn, J., & Lame, I. (2007). Pain assessment in younger and older pain patients: psychometric properties and patient preference of five commonly used measures of pain intensity. *Pain Medicine*, 8, 601-610.

Petersen, R. C., & Negash, S. (2008). Mild cognitive impairment: An overview. *CNS Spectrums*, 13, 45-53.

Polit, D. F., & Beck, C. T. (2004). *Nursing research: Principles and methods* (7th ed.). Philadelphia: Lippincott Williams & Wilkins.

Renn, C. L., & Dorsey, S. G. (2005). The physiology and processing of pain: A review. *AACN Clinical Issues*, 16, 277-290; quiz 413-275.

Rollman, G. B., Lautenbacher, S., & Johnes, K. S. (2000). Sex and gender differences in responses to experimentally induced pain in humans. In R. B. Fillingim (Ed.), *Sex, gender, and pain, progress in pain research and management* (pp. 165–190). Seattle: WA: IASP.

Rosenquist, R. W., & Rosenberg, J. (2003). Postoperative pain guidelines. *Regional Anesthesia and Pain Medicine*, 28, 279-288.

Roykulcharoen, V., & Good, M. (2004). Systematic relaxation to relieve postoperative pain. *Journal of Advanced Nursing*, 48, 140-148.

Rudberg, M. A., Pompei, P., Foreman, M. D., Ross, R. E., & Cassel, C. K. (1997). The natural history of delirium in older hospitalized patients: A syndrome of heterogeneity. *Age Ageing*, 26, 169-174.

- Samabub, S., Petpichetchian, W., & Kitrungrrote, L. (2009). Psychometric properties of pain intensity scales comparing among postoperative adult patients, elderly patients without, and elderly patients with early cognitive impairment. A Proceeding: Research Presentation, Prince of Songkla University, Pattani Campus, Thailand. July 4, 2009.
- Scherder, E. J., & Bouma, A. (2000). Visual analogue scales for pain assessment in Alzheimer's disease. *Gerontology*, *46*, 47-53.
- Serlin, R. C., Mendoza, T. R., Nakamura, Y., Edwards, K. R., & Cleeland, C. S. (1995). When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain*, *61*, 277-284.
- Shen, Q., Sherwood, G. D., McNeill, J. A., & Li, Z. (2008). Postoperative pain management outcome in Chinese inpatients. *Western Journal of Nursing Research*, *30*, 975-990.
- Shi, Q., Wang, X. S., Mendoza, T. R., Pandya, K. J., & Cleeland, C. S. (2009). Assessing persistent cancer pain: A comparison of current pain ratings and pain recalled from the past week. *Journal of Pain Symptom Management*, *37*, 168-174.
- Simons, W., & Malabar, R. (1995). Assessing pain in elderly patients who cannot respond verbally. *Journal of Advanced Nursing*, *22*, 663-669.
- Singchangchai, P., Khampalikit, S., & Na-sae, T. (1996). Using research: Principle and practice. Songkhla, Thailand: Tame Printing.

- Sjostrom, B., Dahlgren, L. O., & Haljamae, H. (1999). Strategies in postoperative pain assessment: validation study. *Intensive & Critical Care Nursing, 15*, 247-258.
- Smith, M. J., Breitbart, W. S., & Platt, M. M. (1995). A critique of instruments and methods to detect, diagnose, and rate delirium. *Journal of Pain and Symptom Management, 10*, 35-77.
- Stolee, P., Hillier, L. M., Esbaugh, J., Bol, N., McKellar, L., & Gauthier, N. (2005). Instruments for the assessment of pain in older persons with cognitive impairment. *Journal of the American Geriatrics Society, 53*, 319-326.
- Taylor, L. J., Harris, J., Epps, C. D., & Herr, K. (2005). Psychometric evaluation of selected pain intensity scales for use with cognitively impaired and cognitively intact older adults. *Rehabilitation Nursing, 30*, 55-61.
- Taylor, L. J., & Herr, K. (2002). Evaluation of the Faces Pain Scale with minority elders. *Journal of Gerontological Nursing, 27*, 15-23.
- Taylor, L. J., & Herr, K. (2003). Pain intensity assessment: A comparison of selected pain intensity scales for use in cognitively intact and cognitively impaired African American older adults. *Pain Management Nursing, 4*, 87-95.
- Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society, 40*, 922-935.
- Tzeng, J. I., Chou, L. F., & Lin, C. C. (2006). Concerns about reporting pain and using analgesics among taiwanese postoperative patients. *The Journal of Pain, 7*, 860-866.

- Unruh, A. M. (1996). Gender variations in clinical pain experience. *Pain, 65*, 123-167.
- Waltz, C. F., Strickland, O. L., & Lenz, E. R. (2005). *Measurement in nursing and health research* (3rd ed.). New York: Springer.
- Wang, X. S., Mendoza, T. R., Gao, S. Z., & Cleeland, C. S. (1996). The Chinese version of the Brief Pain Inventory (BPI-C): Its development and use in a study of cancer pain. *Pain, 67*, 407-416.
- Ward, S. E., Goldberg, N., Miller-McCauley, V., Mueller, C., Nolan, A., Pawlik-Plank, D., et al. (1993). Patient-related barriers to management of cancer pain. *Pain, 52*, 319-324.
- Ware, L. J., Epps, C. D., Herr, K., & Packard, A. (2006). Evaluation of the Revised Faces Pain Scale, Verbal Descriptor Scale, Numeric Rating Scale, and Iowa Pain Thermometer in older minority adults. *Pain Management Nursing, 7*, 117-125.
- Weetman, C., & Allison, W. (2006). Use of epidural analgesia in post-operative pain management. *Nursing Standard, 20*, 54-64; quiz 66.
- Weiner, D., Peterson, B., & Keefe, F. (1999). Chronic pain-associated behaviors in the nursing home: Resident versus caregiver perceptions. *Pain, 80*, 577-588.
- Wong, E. M., & Chan, S. W. (2008). The pain experience and beliefs of Chinese patients who have sustained a traumatic limb fracture. *International Emergency Nursing, 16*, 80-87.
- Xu, G., Meyer, J. S., Huang, Y., Du, F., Chowdhury, M., & Quach, M. (2003). Adapting mini-mental state examination for dementia screening among illiterate or

- minimally educated elderly Chinese. *International Journal of Geriatric Psychiatry*, 18, 609-616.
- Yang, J. Y., Huang, W. Y., Yang, X., Wang, J. H., Lei, M. Y., Zhang, N., et al. (2008). Investigation on cognitive impairment among elders in urban community of Guiyang. *Chinese Journal of Public Health*, 24, 1167-1168 (in Chinese).
- Yin, L., Sun, J.P., Yang, Y.Y., Hu, X. Y. & Chen, Y. (2000). *Geriatrics nursing*. Beijing: People's Health Press (in Chinese).
- Yu, E. S., Liu, W. T., Levy, P., Zhang, M. Y., Katzman, R., Lung, C. T., et al. (1989). Cognitive impairment among elderly adults in Shanghai, China. *Journal of Gerontology*, 44, 97-106.
- Zhang, M. (1991). Comparison of several screening instruments for dementia. *Zhonghua Shen Jing Jing Shen Ke Za Zhi*, 24, 194-196 (in Chinese).
- Zhang, M. Y. (1998). *Neurology assessment book*. Changsha: Hunan technology press (in Chinese).
- Zhang, W. F., Cheng, Y., & Chen, G. (2001). Aging of population and continuable development. *Zhongguo Lao Nian Xue Za Zhi*, 6, 471-475 (in Chinese).
- Zhang, Z. X., Hong, X., Li, H., Zhao, J. H., Huang, J. B., Wei, J. et al. (1999). The Mini-Mental state examination in the Chinese residents population aged 55 years and over in the urban and rural areas of Beijing. *Chinese Journal of Neurology*, 32, 149-153 (in Chinese).
- Zhang, Z. X., Zahner, G. E., Roman, G. C., Liu, J., Hong, Z., Qu, Q. M., et al. (2005).

Dementia subtypes in China: Prevalence in Beijing, Xian, Shanghai, and Chengdu. *Archives of neurology*, 62, 447-453.

Zhao, J. J., Lu, X. Y., Zhao, C. F., & Tuo, Y. (2002). The research on the correlation between the numerical rating scale and verbal rating scale and the improvement on pain assessment. *Xian Dai Hu Li*, 9, 660–661 (in Chinese).

Zhou, Y., Furgang, F. A., & Zhang, Y. (2006). Quality assurance for interventional pain management procedures. *Pain Physician*, 9, 107-114.

APPENDICES

APPENDIX A
INFORMED CONSENT FORM

My name is Yinghua Zhou and I am a master student in the Faculty of Nursing, Prince of Songkla University, Thailand. I am also a nurse in the Surgical Intensive Care Unit, the Second Affiliated Hospital of Kunming Medical University. I am conducting a research study entitled “Psychometric Properties of Pain Intensity Scales Comparing Among Postoperative Adult Patients, Elderly Patients Without and With Mild Cognitive Impairment in China”. It is therefore expected that the findings of this study will contribute to an improvement on postoperative pain management. This study has been approved by the Institutional Review Board of Faculty of Nursing, Prince of Songkla University, Thailand, and is also granted permission by the Second Affiliated Hospital of Kunming Medical University. You are asked to participate in this research project. If you decide to participate in this study voluntarily, I will initiate the following procedures:

1. On the preoperative day, you will be asked about your personal information and health history. For the elderly, you will be also administered the questionnaire to test your cognitive level. Then all of you will be administered the questionnaire to rate your pain intensity level. The whole process may take you about 10 minutes and for the elderly it will take about 25 minutes.
2. During 49-72 hours after surgery, the researcher will interview you

and you will be asked to finish two questionnaires about your pain intensity level and pain interference level. The whole process will take about 20 minutes.

3. During 73-96 hours after surgery, you will be given the questionnaire to rate your pain intensity level again as the preoperative day. At last, you will be asked to choose the most preferred scale and simplest scale based on your opinion. The whole process will take you about 15 minutes.

4. Other than that, you will be given the same routine care throughout your hospital stay.

Risks and discomforts:

There is no evidence shown risk related to finish the questionnaires. However, there is a possibility that some questions asking about your pain level may make you feel physical or psychological discomfort or the questionnaires used in this study will be a burden to you, disturbing your rest after surgery. When the above situations happen to you, please let me know. I will stop the interview and let you take a rest and help you solve the discomfort if you want. Then I can continue the interview when you feel better and are willing to continue. There is no compensation to you for your participation in this study.

Benefits:

The finding of this research will help nurses to enhance the quality of surgical care for patients by offering accurate pain assessment. It will also provide useful information for future research related to this area.

Confidentiality:

All information and your responses in this study will remain confidential, only the researcher, the advisors and the research committee of this study are eligible to access the data. Neither your name nor any identifying information will be used in the report.

Participation and withdrawal:

Your participation in this study is voluntary. Returning the forms given indicates that you understand what is involved and you agree to participate in this study. You have the right to withdraw from the participation at any time.

Lastly, you can contact me by phone 13544579626 if you have any questions or suggestions or cannot participate. If you agree to participate in this study, please sign your name. If you feel uncomfortable to sign but willing to participate, please also let me know. Thank you for your cooperation!

.....
Name of participant	Signature of participant	Date
..Yinghua Zhou.....
Name of researcher	Signature of researcher	Date

APPENDIX B
INSTRUMENTS

APPENDIX B1

DEMOGRAPHIC DATA QUESTIONNAIRE (DDQ)

Preoperatively, the researcher will get the following information from the medical records or by interview.

Date _____ Name _____ Code _____

Ward _____ Bed number _____ Admission time _____

Section One: Demographic Data

1. Age _____ (yrs)
 1. 20-44
 2. 45-59
 3. ≥ 60 without CI
 4. ≥ 60 with mild CI
2. Gender
 1. Male
 2. Female
3. Religion
 1. Buddhism
 2. Islam
 3. Christianity
 4. Others _____
4. Educational level
 1. No education
 2. Primary school

3. () Secondary school

4. () High school

5. () Diploma and bachelor degree or higher

5. Occupation 1. () Government employee

2. () Private employee

3. () Unemployed

4. () Retired

5. () Student

6. () Others _____

6. Marital status 1. () Never married

2. () Married

3. () Divorced/ Separated

4. () Widowed

Section Two: Health Related Information

Date of surgery ____ / ____ / ____

Starting time for surgery _____

Finishing time for surgery _____

Operation sites _____

Previous history of surgery 1. () No

2. () Yes, specify ____ number of times

APPENDIX B2**CHINESE VERSION OF THE MINI-MENTAL STATE EXAMINATION
(CMMSE)**

There are seven dimensions which cover a broad range of cognitive domains in the CMMSE. These seven dimensions include orientation to time (5 points), orientation to place (5 points), registration of three words (3 points), language (8 points), recall of three objects (3 points), attention and calculation (5 points), and visual construction (1 point). The possible scores range from 0 to 30. The questions in the CMMSE are as follows.

CMMSE (Zhang, 1998)

1. What year it is?
2. What season it is?
3. What is today's date?
4. What day of the week it is?
5. What month it is?
6. What province we are in?
7. What city we are in?
8. What district we are in?
9. Which floor we are on?

10. Which place we are in?

11. Please repeat the following three objects] ball, national flag, and tree.

I want you to remember these three objects and I will ask you in a little while to tell me again what these three objects are.

12. Starting from 100, ask the client subtract 7 from the remainder (up to 5 times).

13. Ask the client to tell you what the three objects were that you wanted them to remember.

14. Please tell me what this is? (Show the client a watch)

Please tell me what this is? (Show the client a pencil)

15. Please repeat the following phrase: “Forty-four stone lions”.

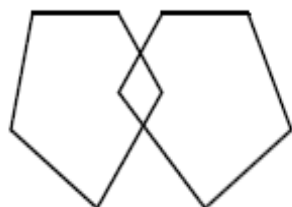
16. Please read this sentence and do what it says: “Close your eyes”.

17. Please take this paper by using your right hand and fold it by using your both hands.

Then, put the paper on your leg.

18. Please say a meaningful sentence.

19. Please copy this design: two overlapping pentagons.



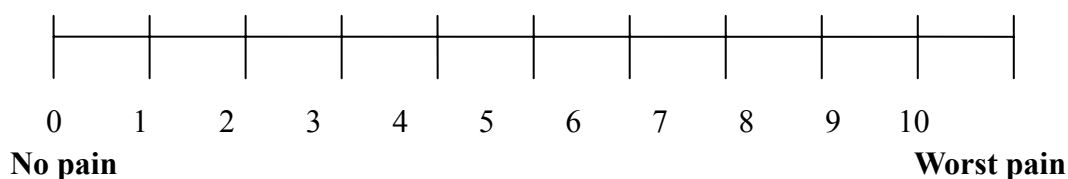
APPENDIX B3

PATIENT'S PAIN INTENSITY ASSESSMENT FORM ONE

Please rate the intensity of the **vividly remembered painful experience** that you had in your life (e.g., toothache, headache, stomachache, childbirth, and back pain) by using the following five pain scales.

▣ Numeric Rating Scale (NRS)

Choose the number that represents your pain intensity **by marking (√) on the number.**



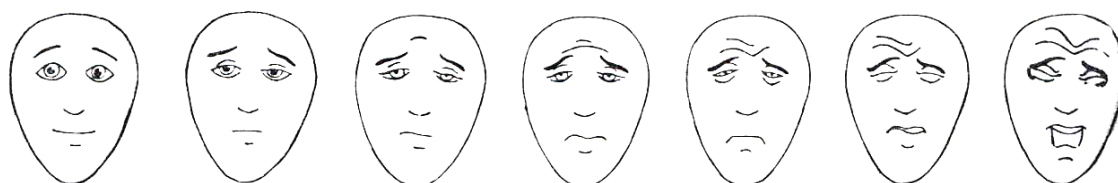
▣ Verbal Descriptor Scale (VDS)

Choose the word that represents your pain intensity **by marking (√) on the word.**

No pain Slight pain Moderate pain Severe pain Unbearable pain

▣ Faces Pain Scale (FPS)

Choose the face that represents your pain intensity **by marking (√) on the face.**



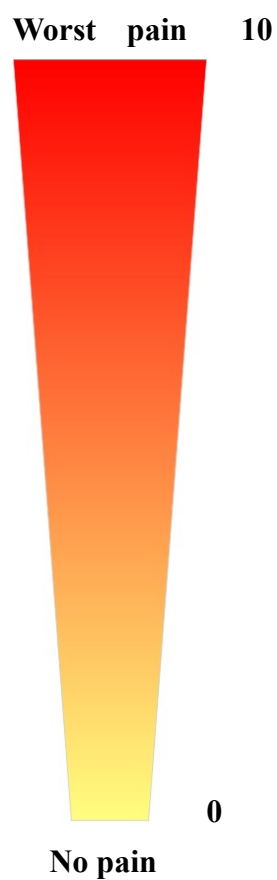
▮ Numerical Box-21 Scale (BS-21)

Choose the box that represents your pain intensity **by marking (✓) in the box.**

0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

No Pain**Worst Pain****▮ Colored Analogue Scale (CAS)**

Choose **the position** that represents your pain intensity **by marking a horizontal line on the scale.**

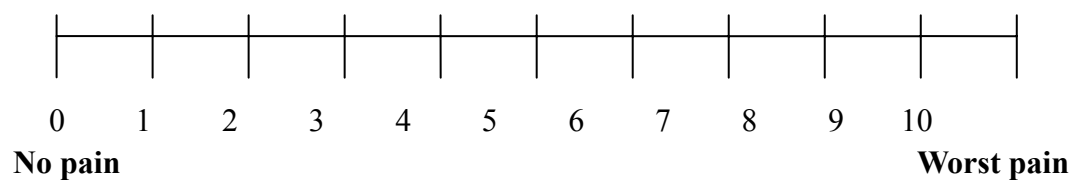


APPENDIX B 4**PATIENT'S PAIN INTENSITY ASSESSMENT FORM TWO**

Rating the intensity of your **current pain** and **daily retrospective worst, least, and average pain that you had in the past 24 hours** by using the following five pain scales.

□ Numeric Rating Scale (NRS)

Choose the number that represents your pain intensity by marking (✓) on the number.



Current pain intensity _____

Worst pain intensity during the past 24 hours _____

Least pain intensity during the past 24 hours _____

Average pain intensity during the past 24 hours _____

▮ Verbal Descriptor Scale (VDS)

Choose the word that represents your pain intensity by marking (√) on the word.

No pain Slight pain Moderate pain Severe pain Unbearable pain

Current pain intensity _____

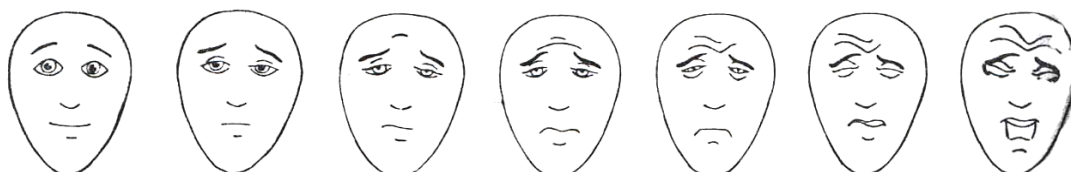
Worst pain intensity **during the past 24 hours** _____

Least pain intensity **during the past 24 hours** _____

Average pain intensity **during the past 24 hours** _____

□ Faces Pain Scale (FPS)

Choose the face that represents your pain intensity by marking (✓) on the face.



Current pain intensity _____

Worst pain intensity during the past 24 hours _____

Least pain intensity during the past 24 hours _____

Average pain intensity during the past 24 hours _____

▣ Numerical Box-21 Scale (BS-21)

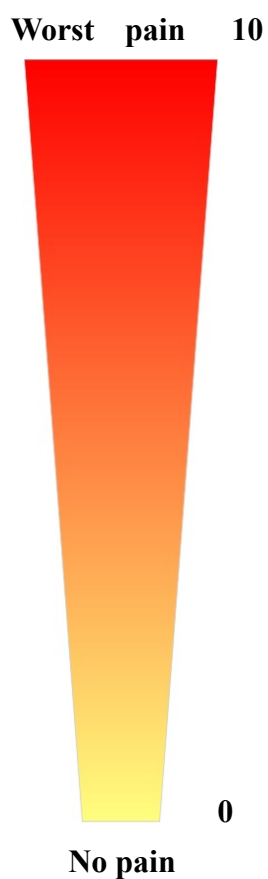
Choose the box that represents your pain intensity by marking (√) in the box.

0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

No Pain**Worst Pain****Current pain intensity** _____**Worst pain intensity during the past 24 hours** _____**Least pain intensity during the past 24 hours** _____**Average pain intensity during the past 24 hours** _____

▮ Colored Analogue Scale (CAS)

Choose the position that represents your pain intensity by marking a horizontal line on the scale.



Current pain intensity _____

Worst pain intensity during the past 24 hours _____

Least pain intensity during the past 24 hours _____

Average pain intensity during the past 24 hours _____

APPENDIX B 5

The MODIFIED PAIN INTERFERENCE SCALE (PIS)

Marking (√) on the number that represents the degree to which the
worst pain during the past 24 hours has interfered with your

A. General Activity

0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

B. Mood

0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

C. Walking Ability

0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

D. Relationships with other people

0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

E. Sleep

0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

F. Enjoyment of life

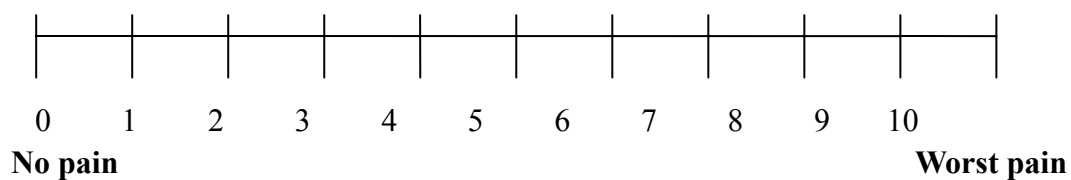
0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

G. Ability to think and make decision

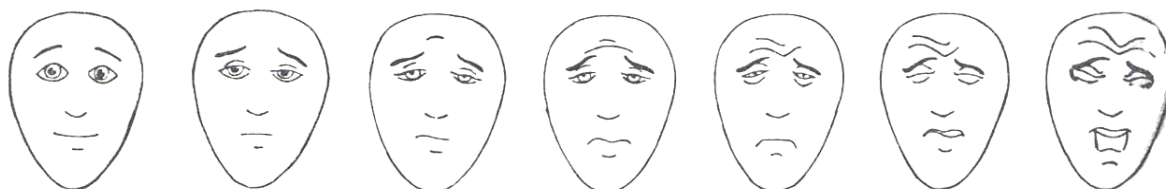
0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

APPENDIX B 6**SCALE PREFERENCE QUESTIONNAIRE (SPQ)**

Based on your opinion, please **rank the following scales** from the most preferred to the least preferred *with 1 = most preferred and 5 = least preferred*.

___ Numerical Rating Scale (NRS)**___ Verbal Descriptor Scale (VDS)**

No pain Slight pain Moderate pain Severe pain unbearable pain

___ Faces Pain Scale (FPS)

____ **Numerical Box-21 Scale (BS-21)**

0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

No pain

Worst pain

____ **Colored Analogue Scale (CAS)**

Worst pain 10



0

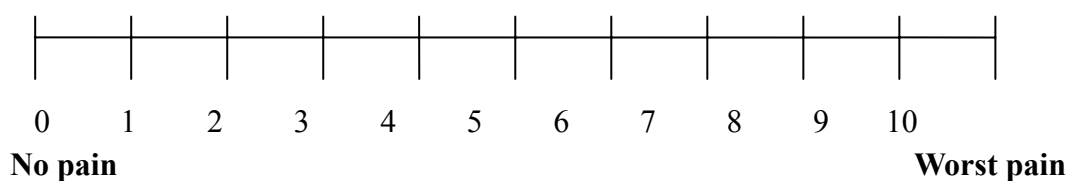
No pain

APPENDIX B 7

SCALE SIMPLICITY QUESTIONNAIRE (SSQ)

Based on your opinion, please **rank the following scales** from the simplest to the least simple with *1 = simplest and 5 = least simple*.

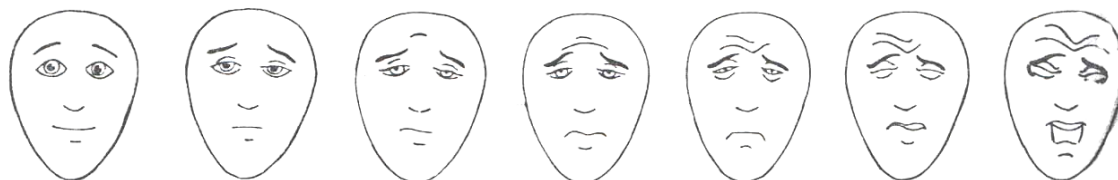
___ Numerical Rating Scale (NRS)



___ Verbal Descriptor Scale (VDS)

No pain Slight pain Moderate pain Severe pain unbearable pain

___ Faces Pain Scale (FPS)



____ **Numerical Box-21 Scale (BS-21)**

0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

No pain

Worst pain

____ **Colored Analogue Scale (CAS)**

Worst pain 10



0

No pain

APPENDIX B 8

SCALE ACCURACY CHECKLIST (SAC)

Type of errors	VDS	NRS	FPS	CAS	BS-21
	n (%)	n (%)	n(%)	n (%)	n (%)
Ratings outside the scale range (ratings between scale units or above/below the end-points of the scale)					
No rating on the scale					
More than one rating on a single scale					
A range of ratings on one scale					
Response falling between two numbers, words, or facial expressions					
Mistake in ordinal understanding of the scale					

APPENDIX C

TABLES

Table 14
Amount of Errors of Each Scale Classified by Six Different Types of Errors in Four Groups (N = 200)

Type of errors	VDS	NRS	FPS	CAS	BS-21
	n (%)	n (%)	n (%)	n (%)	n (%)
Ratings outside the scale range (ratings between scale units or above/below the end-points of the scale)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
No rating on the scale	0 (0)	0 (0)	0 (0)	9 (4.5)	2 (1)
More than one rating on a single scale	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
A range of ratings on one scale	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Response falling between two numbers, words, or facial expressions	0 (0)	1 (0.5)	0 (0)	0 (0)	0 (0)
Mistake in ordinal understanding of the scale	0 (0)	2 (1)	1 (0.5)	0 (0)	0 (0)

Table 15

Detail Statistics of F-test: Test the Differences of Test-retest Reliability Coefficients Comparing Among Different Age Groups and Levels of CI, Using one-way ANOVA (N = 153, n = 43 for Young adults, n = 49 for Middle-aged Adults, n = 40 for Elderly Without CI, and n = 31 for Elderly With Mild CI)

	Four groups	SS	df	MS	F	p-value
VDS	Between groups	0.68	3	0.28	0.35	.79
	Within groups	104.45	161	0.65		
NRS	Between groups	1.18	3	0.39	0.52	.67
	Within groups	119.36	159	0.75		
FPS	Between groups	3.06	3	1.02	1.30	.28
	Within groups	126.45	161	0.79		
CAS	Between groups	6.38	3	2.13	2.43	.07
	Within groups	133.33	152	0.88		
BS-21	Between groups	2.97	3	0.99	1.23	.30
	Within groups	128.10	159	0.81		

APPENDIX D
METHODS IN DATA ANALYSIS

In this study, One-way ANOVA or Kruskal-Wallis test were used to test the differences of concurrent validity, convergent validity, and test-retest reliability of each scale among the four groups. The principle to conduct these analyses was based on the following formula (Nunnally & Beirnstein, 1994):

$$r_{xy} = \frac{\sum Z_x Z_y}{N}$$

r_{xy} = the correlation between x and y

Z_x = the standardized score of x

Z_y = the standardized score of y

$\sum Z_x Z_y$ = the sum of $Z_x Z_y$

N = the total number of subjects

Therefore, the procedures to conduct the One-way ANOVA or Kruskal-Wallis test were as follows. Firstly, the pain scores for each patient were transformed into standardized score: Z_x , Z_y (e.g. $Z_x = Z_{\text{pretest}}$, $Z_y = Z_{\text{posttest}}$). Secondly, the $Z_x Z_y$ for each patient was calculated. Finally, the researcher used the $Z_x Z_y$ of each patient to conduct One-way ANOVA or Kruskal-Wallis test.

APPENDIX E

LIST OF EXPERT PARTICIPANTS

Three Experts in Examining the Content Validity of the Scale Accuracy

Checklist (SAC) and the Modified Pain Interference Scale (PIS):

1▣ Assoc. Prof. Dr. Praneed Songwathana, Ph.D., RN

Faculty of Nursing, Prince of Songkla University, Thailand.

2▣ Assist. Prof. Dr. Sasitorn Phumdoung, Ph.D., RN

Faculty of Nursing, Prince of Songkla University, Thailand.

3▣ Assist. Prof. Dr. Sasikarn Nimmanratch, Ph.D.

Department of Anesthesiology, Faculty of Medicine, Prince of Songkla University,
Thailand.

One Technician Expert in Modifying the Picture of the Colored Analog

Scale: Mr. Withoon Sangkharak, Audio-visual Aids (AVA) technician, Faculty of
Nursing, Prince of Songkla University, Thailand.

VTTAE

Name Miss Yinghua Zhou

Student ID 5110420047

Educational Attainment

Degree	Name of Institution	Year of Graduation
Bachelor in Nursing Science	Kunming Medical University, China	2006

Scholarship Awards during Enrollment

MOU scholarship, Prince of Songkla University

Teaching Assistant Scholarship, Graduate School, Prince of Songkla University

Work-Position and Address

Nurse, Surgical Intensive Care Unit, the Second Affiliated Hospital of Kunming Medical University, 650100, China.

E-mail: zyh7069813@hotmail.com