Evaluating the psychometric properties and the clinical feasibility of a Chinese version of the Doloplus-2 scale among cognitively impaired older people with communication difficulty.

Introduction

Adequate pain assessment is a critical part of effective pain management in long-term care facilities, implicated with quality of care and quality of life for older people with dementia (Stolee *et al.* 2005). Although self-report of pain is considered the "gold standard" for pain assessment, those with moderate to severe dementia may have difficulty communicating their pain or even reduced awareness of pain. To solve the dilemma of pain assessment, behaviorally observed methods have been suggested for use with this group, and several behavioural pain assessment tools have been developed. However, most instruments published in English lacked satisfactory reliability, validity and clinical usefulness (Herr *et al.* 2006, Smith *et al.* 2005, Stolee *et al.* 2005).

A systematic review of behavioural pain assessment tools not just focused on those publications in English was conducted by Zwakhalen *et al* (2006), and based on quality judgment criteria such as the psychometrics properties, sensitivity and feasibility. The findings demonstrated the Doloplus-2 scale was one of the most appropriate scales currently available for assessing pain in older people with severe dementia. Many key indicators of pain in cognitively impaired older people, noted in the literature and by the American Geriatrics Society (AGS) panel on persistent pain in older persons, are comprised comprehensively in the Doloplus-2 scale (AGS panel 2002, Herr *et al* 2006). Additionally, the distinguishing characteristic of the tool is the options of items closely relevant to observe subtle changes in usual behavioral expressions in different situations, which truly reflect nursing staffs' view about how to recognize pain in older people with advanced stage of dementia (Blomqvist & Hallberg 2001, AGS panel 2002).

Background

The Doloplus scale, the precursor of Doloplus-2 scale was developed by Wary *et al* in 1993 for assessing pain in older people and inspired by the Douleur Enfant Gustave Roussy (DEGR) scale for young children (Lafebvre-Chapiro & DOLOPLUS group. 2001). Subsequently, a DOLOPLUS group was formed by geriatricians in France to refine the Doloplus scale into the present version of Doloplus-2 scale mainly for evaluating pain in older people with communication problems (Lafebvre-Chapiro & DOLOPLUS group. 2001, Holen *et al.* 2007). The Doloplus-2

scale involves ten main types of pain behaviors in cognitively impaired older people, categorized into three subscales including somatic reactions (somatic complaints, protective body postures adopted at rest, self-protection of sore areas, facial expression and sleep pattern), psychomotor reactions (washing and/or dressing, physical activity) and psychosocial reactions (verbal/vocal communication daily, social life and behavioral problems). Ten items are described separately by four levels of behavioral expressions from zero to three representing increasing intensity of pain. A score greater than or equal to five out of 30 is indicative of pain (Lafebvre-Chapiro & DOLOPLUS group. 2001).

The reliability and validity of the French version of Doloplus-2 has been evidenced in diverse people and settings (Lafebvre-Chapiro & DOLOPLUS group. 2001). According to the DOLOPLUS team, the French version of Doloplus-2 was tested with more than 500 older people in a early study, showing satisfactory internal consistency and concurrent validity between Doloplus-2 and a visual analog scale (VAS) (Lafebvre-Chapiro & DOLOPLUS group. 2001). In Pautex *et al*'s study (2005), the French Doloplus-2 was used to examine the validity of four pain self-assessments among hospitalized older people with dementia, and the results demonstrated that the tool correlated moderately with these self-report scales. Moreover, Pautex *et al* (2007) adopted VAS to validate this tool in hospitalized older people with and without dementia who can use VAS reliably. The findings showed Doloplus-2 scale could predict 41% variance of VAS and had adequate internal consistency.

In addition, the French version of Doloplus-2 scale was further translated into Norwegian and Dutch for validation, and adequate psychometric qualities were supported in these studies (Bjoro *et al.* 2008, Holon *et al.* 2005, Zwakhalen *et al.* 2006). Holen *et al*'s pilot study (2005) used expert pain rating as the criterion to validate the Norwegian version of Doloplus-2 (N-Doloplus-2) in cognitively impaired patients unable to self-report pain; reasonable criterion validity and clinical feasibility were reported. Good inter-rater reliability was also reported in Holen *et al'*s later study (2007). In addition, Zwakhalen *et al* (2006) evaluated the pshychometric quality of Dutch version of three pain behaviorally observed tools among nursing home residents with dementia, and the

results demonstrated the Doloplus-2 had adequacy internal consistent (Cronbach's α=0.74-0.75) and acceptable congruent validity (r=0.29-0.34).

Although several researchers indicated the performance of Doloplus-2 scale required more training which may impede clinical feasibility, especially the assessment of the psychosocial items (Zwakhalen *et al* 2006, Holen *et al.* 2007), the clinical value of the tool for detecting pain in older people with advanced dementia is undoubted. Moreover, no matter which pain assessment scale is being used, it remains a crucial task for clinical practitioners to observe and interpret pain behaviors in this group. In the past decade, there has been a growing public awareness and studies related to pain of older people with dementia. Until now, no adequate behaviorally observed scales are available for measuring pain among older people with dementia in Taiwan. In response to this issue, to use the previously developed French version of Doloplus-2 scale with promising psychometric qualities for cross-culture validation seemed possible, instead of developing a new scale with unknown measurement properties (Cha *et al.* 2007). Nevertheless, the quality of translation and appropriate methods of evaluating the psychometric properties of translated instrument should be ensured (Maneesriwongul & Dixon 2004).

The study

The major purpose of this study was to translate the French version of Doloplus-2 (F-Doloplus-2) into Chinese (C-Doloplus-2) and further evaluate the reliability and validity of the C-Doloplus-2 among cognitively impaired older people with communication difficulty in Taiwan. Specific aims were as follows:

(1) to evaluate the internal consistency and inter-rater reliability of C-Doloplus-2.

(2) to validate the construct validity through examining the known association between C-Doloplus-2 and empirically supported correlates of pain such as the past pain history, the number of falls during the recent six months, disability, agitation and depression.

(3) to validate the construct validity through the Principal Factor Analysis (PFA).

cipal Factor Analysis (PFA).

l Factor Analysis (PFA).

ctor Analysis (PFA).

Analysis (PFA).

lysis (PFA).

s (PFA).

FA).

(4) to evaluate the clinical feasibility of the C-Doloplus-2.

Methodology

Instruction translation and validation was carried out through a three-phase process in Taiwan from September 2007 to May 2008.

Phase I. Forward-translation and back-translation

The original F-Doloplus-2 was translated into Chinese by a bilingually licensed translator, working in the translation agency for 5 years. After forward-translation was completed, the second bilingual person, a specialist in multiple languages, living in France for 8 years, performed the blind back-translation into French. Finally, a native speaker of French, teaching French in a language center, assisted in comparing the linguistic difference between the original instrument and back-translated version. Two items were judged as discrepant and were re-checked by the prior two bilingual persons together until the translation was satisfactory. The third option of item 4 ‘staring’ was revised into ‘dull eyes’. The item 7 ‘mobility’ was revised into ‘physical activity’. Additionally, back-translation of French instruction of Doloplus-2 was also completed in the same procedures.

Although a back-translation procedure can produce linguistic equivalence of two versions, whether the translated instrument measure the same domain across culture cannot be guaranteed (Hyrkas *et al.* 2003). To assure each item measured pain behaviours of older people with dementia and matching the Taiwanese culture, 5 experts, including 2 gerontological nursing educators, 2 neurological physicians and a gerontological head nurse were invited to examine the content of C-Doloplus-2 and rated each item on a 4-point Likert scale from relevant (4) to irrelevant (1).

Then the content validity of index (CVI) was calculated by summing the percentage agreement of all items that were given a rating of '3' or '4' (Lin *et al.* 2007), yielding a perfect agreement of 100%. Only the option 'insomnia, affecting morning waking time' of item 5 'sleep pattern' is recommended to rephrase to 'insomnia, affecting morning waking time or mental state during awake period'. Because older people residing in some institutions in Taiwan are enforced to wake up at the same time, just relying on the morning waking time makes it hard to judge whether sleep pattern of older people is affected by pain.

Phase II. Pilot testing

To examine the initial psychometric characteristics of the translated version and ensure no potential problems existed in instrument or performance, a pretest procedure was implemented in a sample of 36 older people with dementia, chosen from a veteran dementia special care unit in Taipei County. One registered nurse (RN) and 4 nursing assistants (NAs) in charge of these residents received training for 1 hour to know how to perform the C-Doloplus-2. Then the C-Doloplus-2 was completed by RN, RAs and one researcher simultaneously for assessing the present pain of residents. Average age of subjects was 82.42 (SD 5.10; range 73-95), and the majority were uneducated (n=16, 44.4%). As to residents' cognitive status and activities of daily living, the mean Mini-Mental State Exam (MMSE) score was 11.36±5.61 and the mean Barthel index score was 81.39±20.09. In addition, the clinical feasibility of the C-Doloplus-2 was investigated by a brief question "Do you think the C-Doloplus-2 is appropriate for assessing pain in cognitively impaired older people with communication difficulty". A 5-point Likert-type scale, which ranged from 5 (strongly agree) to 1 (strongly disagree), was rated by each administrator. Subsequently, these nursing staff were also encouraged to write down comments about the C-Doloplus-2 for future improvement.

The reliability of the C-Doloplus-2 scores assessed by RN, RAs and researcher ranged form 0.77-0.84. The intraclass correlation coefficient (ICC) of the C-Doloplus-2 scores between RN and researcher was 0.85. The average score of clinical feasibility assessed by 5 nursing staff was

3.8 (SD 0.45; range 3-4). All staff agreed that 10-items of C-Doloplus-2 can represent vital pain behavioral indicators of older people with dementia and these items are easy to comprehend. However, two staff expressed that psychosocial items are difficult to observe in older people with severe dementia in clinical practice, recommending the need for a more detailed instruction for future performance. Based on nursing staff response, a more detail user manual, expended from the original instruction of C-Doloplus-2, was developed for institutional use. In the user manual, the standardized procedure for performance was addressed step by step and each item was augmented by the definition of terminology, reference indicators, scoring method, assessment method, and practice exercises.

Phase III. Validation the psychometric qualities of C-Doloplus-2

The main validation procedures were performed in a prospective study in five dementia special care units in the Northern Taiwan. The internal consistency and inter-rater reliability between RNs and trained research assistants (RAs) were evaluated. In addition, exploratory factor analysis and the construct validity were also evaluated through examining the association between pain and known correlates of pain such as the pain history, presence of pain related conditions, functional disability, agitation and depression based on Snow *et al*'s (2004)the Know Correlates Validity Model.

Participants

Residents with moderate to severe dementia and RNs were recruited. The inclusive criteria for older people were: (1) being older than 65 years; (2) living in the dementia special care unit; (3) having the diagnosis of dementia; (4) present MMSE scores under 18. The inclusive criteria for RNs were (1) having worked in this dementia special care unit at least one month before data collection; (2) in charge of the resident. In addition, the research assistants (RAs) employed to observe the pain of residents received intensive training. During training, the ICC for the C-Doloplus-2 scale between researcher and RAs ranged from 0.91-0.93.

Measurements

The characteristics of older people with dementia included demographic data, medical

conditions, the past pain history and the numbers of falls during the recent six months were collected by a data sheet. The dementia severity of older people was examined by the MMSE.

Measurement of mental status. The mental status of residents was measured by the Mini-Mental State Examination (MMSE) (Folstein *et al*. 1975) which is an 11-item and 30-point instrument. The lower scores represent higher levels of cognitive impairment. The commonly recommended cut-off scores for normal, mild, moderate and severe cognitive impairment, respectively, are 23/24, 17/18 and 11/10 (Juva *et al*. 1994, Tombaugh & McIntyre 1992). The test-retest reliability raged from 0.83-0.99 and the criterion-related validity of MMSE and the Wechsler adult intelligence scale (WAIS) was supported (r= 0.66-0.78) (Folstein *et al*. 1975). The internal consistency of the MMSE has been shown to be excellent in Taiwan (α=0.91) (Yip *et al.* 1992). In addition, pain of residents was measured by the C-Doloplus-2 scale, disability, agitated behavior and depression were also measured by different with promising psychometric quality.

Measurement of disability. The functional disability of residents was measured using the Barthel index. It is a 10-item and 100-point clinical rating scale. Lower scores represent higher levels of physical dependence (Mahoney & Barthel 1965). The psychometric properties of the Barthel index have been supported in many studies (Green, & Young 2001, Sainsbury *et al.* 2005). The Barthel index has been tested widely in Taiwan, and the internal consistency was 0.89-0.92 and the inter-rater reliability was 0.94. Good concurrent validity was reported by examining the correlations the Barthel index and Fugl-Meyer motor assessment (FM) (r= 0.78-0.80) and Berg balance scale (BBS) (r=0.89-0.94) (Hsueh *et al*. 2001).

Measurement of agitated behaviours. Agitation of residents was assessed using the Cohen-Mansfield agitation inventory (CMAI). The content consisting of 29 items is administered by caregivers to recall the frequency of those agitated behaviors manifested in the past week and rate the average intensity of these items on a seven-point scale ranging from 1 = never - 7 = several times an hour (Cohen-Mansfield, 1986). Higher scores mean higher severity of agitated behaviors; the internal consistency ranges from 0.63-0.82 and the inter-rater agreement rate ranges from 0.88-0.92. Factor analysis extracted three factors: aggressive behavior, physically non-aggressive

behavior and verbally agitated behavior, which can explain the 37.9% variance (Cohen-Mansfield *et al*. 1989, Miller *et al*. 1995). In Taiwan, a study conducted by Lin *et al*. (2007) translated the original CMAI into Chinese; the semantic equivalence between E-CMAI and C-CMAI (ICC=0.69-0.74), and the technical equivalence of C-CMAI collected by two different methods (ICC=0.63-0.86) have been evidenced.

Measurement of depression. Depressive symptoms of residents were observed by the Cornell scale for depression in dementia (Alexopoulos *et al*. 1988). This is a 19-item clinical observation scale composed of five subscales, including mood-related signs, behavioral disturbances, physical signs, cyclic functions and ideational disturbance. The frequency of depressive symptom in the past week is rated by caregivers on a four point scale: 0 = absent, 1 = mild or intermittent, 3 = severe and a = unable to evaluate. Higher scores reflect a higher level of depression. A cut-off score of 6 is commonly used to distinguish those who are highly depressive symptoms (Lam *et al*. 2004). The internal consistency was 0.84 and the weighted kappa for inter-rater reliability was 0.67. The concurrent validity of the Cornell scale and research diagnostic criteria (RDC) was supported (r = 0.83) (Alexopoulos *et al*. 1988). In our pilot study, the alpha coefficient was 0.81 and the ICC for inter-rater reliability was 0.91.

Data collection

After obtaining permission to perform the study, RNs in each institution were told about the study purpose by the researcher and received intensive training about the use of the C-Dloplus-2, following the user manual. For safeguarding reliability of behavioral observations, one of researchers supervised their use of the C-Doloplus-2 and assisted with instruction about practical problems. To ensure these trained RAs familiarity with residents, they were asked to observe and record resident's painful conditions at rest and after pain-provoked motion each day from 8:00-17:00 for one week. In the last day of the week, before shift change, the C-Doloplus-2 was completed by RNs and RAs simultaneously. The 5-point Likert scale (the same as pilot testing) was also used for scoring the clinical feasibility of C-Doloplus-2. During data collection, based on the medical chart and the interview with nursing staff, residents' demographic information, medical

conditions, the past pain history and the numbers of falls during the recent six months were gathered, and the MMSE, Barthel index, CMAI and the Cornell scale for all residents also administered by RAs.

Data analysis

Statistical analyses were performed by SPSS for windows 11.5. Descriptive statistics were generated for the characteristics of subjects and the distribution of the C-Doloplus-2 mean score. Internal consistencyof the C-Doloplus-2 scales and subscales were determined by Cronbach's alpha, and the intra-rater reliability was examined by ICC. The associations between C-Doloplus-2 and correlates of pain were examined using Pearson's correlation coefficient. Then the PFA with varimax rotation was performed for examining the underlying factor structure. The Eigenvalues and scree plot were used to determine the number of components, and only factors with an Eigenvalue larger than 1 should be extracted (Field 2005). If the factor loading was under 0.3, the item was deleted.

Ethical considerations

The research was approved by the ethical committee at the National Yang-Ming University. Authorisation was obtained from DOLOPLUS group to translate and use the F-Doloplus-2. All participants were informed about the research purpose and asked to sign informed consent, for those people unable to comprehend the information, a legal surrogate consent was acquired. The confidentiality of information and anonymity were assured by researchers.

Results

Sample characteristics

Demographic characteristics of participants were presented in Table 1. There were 241 older people with moderate and severe dementia and 14 RNs in the study. These older people, including 118 female and 123 male, had a mean age of 79.28 years and the mean length of stay in the institution was 38.32 monhs. In terms of older people's medical conditions, the average number of active medical diagnoses was 2.23 (SD 1.51), and the average number of prescribed medications was 6.47 (SD 3.49). Sixty-eight percent of older people had experienced pain. The average

number of falls during the recent six months was 0.50 (SD 1.75). More than half of the older people has severe dementia and were very low in physical functioning. The mean age of RNs was 35.29 years (SD 8.17), with an average 16.83 (SD 19.10) years working experience, and all of them were female. RNs' average time of working in the institution was 4.09 (SD 3.22) years, and the average working time per week was 44 (SD 3.92) hours.

Reliability of the C-Doloplus-2 scale

The means, standard deviations and reliability of three subscales and the overall C-Doloplus-2 were presented in Table 2. The mean C-Doloplus-2 score was 4.49 (SD 4.10) with range from 0-21. Following the cut-off score guideline, ninety-six older people (39.83%) had a score of C-Doloplus-2 equal or greater than 5, signifying pain. Except for the 'psychosocial reactions' subscale, the standardized Conbach's alpha coefficients for the other subscales and the overall C-Doloplus-2 all surpassed the criterion of 0.70. The alpha coefficients did not increase (range, 0.70-0.73) when any of the items were deleted, supporting that the C-Doloplus-2 is a reliable instrument with little measurement error (Firris & Norton 1992). The ICC of the overall C-Doloplus-2 scale between RNs and RAs was 0.81, and of the three subscales ranged from 0.60-0.81.

Validity of the C-Doloplus-2 scale

The Pearson's correlations between C-Doloplis-2 and these known correlates of pain are shown in Table 3. In older people with moderate dementia, only disability was correlated significantly with the C-Doloplus-2 score. However, in older people with severe dementia, both disability and deprssion correlated significantly with the C-Doloplus-2 score, partially confirming the hypothesis that.

The result of the exploratory factor analysis is shown in Table 4. The KMO value for the 10 items scale was 0.71, supporting the use of factor analysis in this sample (Field 2005). Based on the Eigenvalue and scree plot, three factors were extracted, accounting for 65% of the total variance of the C-Doloplus-2. The first factor consisted of all five items of the 'somatic reactions' subscale, and explained 27.43 % of the variance. The second factor consisted of all three items of the

'psychosocial reactions' subscale, and explained 19.86 % of the variance, while the third factor included all two items of the 'psychomotor reactions' subscale, accounting for an additional 19.99% of the variance. Finally, the item-total and item-subtotal correlation was analyzed to examine whether these items were correlated strongly with the subscale and overall scale of the translated instrument, as shown in Table 5. Each item was correlated strongly with the originally belonged subscale, ranged from 0.63 to 0.94.

Clinical feasibility of the C-Doloplus-2 scale

The average score of clinical feasibility assessed by 14 RNs was 4.14 (SD 0.77; range 3-5), supporting the clinical usefulness of C-Doloplus-2 scale. Although several RNs indicated it is difficult to distinguish whether there are behavioral changes in sleep pattern, communication and social life of older people with the advanced stage of dementia in a persistent vegetative state, most RNs agreed the C-Doloplus-2 scale has clinical potential to detect pain in this group

Discussion

This study was the first to validate a behaviorally observed instrument using informant rating method among cognitively impaired older people in Taiwan. The mean scores of the total C-Doloplus-2 demonstrated that older people with moderate and severe dementia had low levels of behavioral expressions of pain. The data from a past survey showed the majority of older people residing in the institution suffer from chronic pain in stead of acute pain; therefore behavioral response to chronic pain is less obvious than acute pain (Weiner *et al.* 1999). The C-Doloplus-2 has an adequate internal consistency for the total scale and almost all subscales. Although, in Zwakhalen *et al*'s (2006) validation study, they also recruited older people with mild dementia as subjects, similar internal consistency scores for the total D-Doloplus-2 scale (alpha 0.75), and for subscales (alpha range 0.63-0.80) were reported. However, both in our study and in Zwakhalen *et al*'s study (2006), the lowest internal consistency was found for the 'psychosocial reactions' subscale. These findings with Cronbach's alpha meeting the criteria of 0.7 are acceptable for a newly-developed scale (LoBiondo-Wood & Haber 1998). Presumably, although slightly poor internal consistency of the 'psychosocial reactions' subscale may be due to the heterogeneity of

items (communication, social life and behavioral problems), the error in measurement may also result from the ambiguous definition of the psychosocial behaviors in the institutions. Abnormal social reactions are related to several possible causes such as dementia severity, delirium and depression (Hadjistavropoulos et al. 2008), therefore, institutional caregivers may have difficulty in recognizing these behavioral manifestations as pain indicators.

Several researchers indicated that the intra-rater and inter-rater reliability of the scale has not yet been adequately established (Bjoro *et al* 2008, Zwakhalen *et al* 2006). Compared with Holen *et al*'s study (2007) including 15 patients to evaluate the inter-rater reliability of the N-Doloplis-2, the inter-rater reliability of the C-Doloplus-2 was established to be more reliable in this study with a larger sample size. The ICC for the total scale was higher than the finding of Holen *et al* (2007), and then the good concordance for the three subscales also further confirmed the inter-rater reliability of the C-Doloplus-2 (Landis & Koch 1977). As we expected, ICC values higher than 0.80 for the inter-rater reliability may be difficult to reach when testing a behaviorally observed instrument for pain among older people with dementia. Due to of the complexity of the language of pain itself, some behaviors of older people with dementia are not typically indicative of pain and the behavioral expressions of pain may mean different things for different assessors (Weiner et al 1999).

Among older people with moderate and severe dementia, a gold standard for pain self-report data could be difficult to obtain. Furthermore, whether self-report of pain assesses the same parameters as behaviorally observed scales of pain still has to be resolved. Therefore, adopting cognitively impaired older people's self-report as a pain criterion to test the criterion validity or the convergent validity of the Doloplus-2 scale may be inappropriate (Pautex *et al.* 2007). A more reliable approach to validate the C-Doloplus-2 scale is to use empirically supported correlates of pain as evidence of the construct validity. However, in our findings, the construct validity of the C-Doloplus-2 scale was only partially supported. Although the association between pain and disability or depression was demonstrated, there was a non-significant association between the pain latent variable and other variables such as the past pain history, numbers of falls and agitation.

Several reasons were speculated to explain the results. First, pain related information under-assessment and under-documentation occurred commonly in the long-term care facilities, therefore it is reasonable to believe that the past pain history collected by research assistants from the medical chart and nursing records could be underestimated. Although the interview with nursing staff was used as a complementary strategy to collect information, recall bias could not be excluded. Second, for older people, although falls may cause some serious consequence such as fractures and head injury, these painful conditions do not accompany every fall. The numbers of falls with physical injury may be a more appropriate indicator for current painful conditions in a validation study. Third, most institutional older people suffered from chronic pain, and agitated behaviors are presumed to be triggered mainly by acute pain rather than chronic pain. Therefore, no significant correlation between pain and agitation may result from less agitated behaviors expressed by our participamts.

The concept of pain for the Doloplus-2 scale is defined as a multidimensional construct with somatic, psychomotor and psychosocial dimensions and, within these domains, observable behaviors and signs may implicate the presence of pain. In the current study, the three factors were identified through the PFA, perfectly congruent with the original three-dimensional structure and the significant item-subtotal correlations between each item and factor scores further provided additional support for the item clusters as subscales of the C-Doloplus-2. Based on our findings, three factors were considered equally important for detecting pain of this group. However, in Pautex *et al*'s study (2007), although the criterion validity was confirmed, the validity of the psychosocial domain was questioned, and the psychosocial items were suggested to remove from the Doloplus-2 scale (Pautex *et al* 2007).

While responding to the above-mentioned incongruity conclusion, a methodological issue should be proposed for further dscussion. Pautex *et al*'s study (2007) used multiple linear regression analysis to investigate what percentage of variability in the VAS, reported by older people with dementia, can be predicted by the Doloplus-2. Due to the main variance of VAS being explained by the somatic dimension, the authors proposed to abbreviate the Doloplus-2 to a 5-item

version; nevertheless, the suggestion seems arbitrary. Undoubtely, the somatic subscale of the Doloplus-2 comprises the 'somatic complaints' item, therefore to be nature and right, the somatic dimension could predict more variance of the pain self-report scale. It is noteworthy to indicate that, athough there are many different methods for evaluating the validity of behaviorally observed pain instruments, not all strategies are appropriate to be used with cognitively impaired older people with communication difficulty.

Finally, regarding the clinical feasibility, although the psychosocial items were mentioned by RNs as almost impossible to perform with older people with severe dementia in a persistent vegetative state, we suggested the psychosocial dimensions should be retained. Based on the concept basis of the Doloplus-2, it was developed for multidimensional assessment of pain in older people with communication difficulty; therefore, the items cover all obvious and less obvious pain-related cues. To avoid the possibility of underassessment and undertreatment of pain, such a comprehensive behaviorally observed instrument should be used in clinical practice (Hadjistavropoulos *et al.* 2008). Although a sensitive scale may decrease the specificity of a behavioral instrument (Bjoro *et al.* 2008), when applied to detecting pain in this group, sensitivity may be more important than specificity.

Limitation

There are several limitations in our study. First, during the process of translation, it was difficult to find another bilingual person to repeat the process of back translation of the discrepant item. Future cross-culture validation studies should use more bilingual persons to perform the translation technique. Second, pain related conditions such as the past pain history and number of falls during the recent six months is usually under-documented in this group. Additional studies are required to adopt other current painful conditions to validate the translated scale. Third, we acknowledge that the observation and score of subtle changes in these pain behaviors may be quite abstract for clinical practitioners. Although all RNs received intensive training about how to perform the C-Doloplus-2 before ratings, it cannot guarantee that variance associated with observers is excluded completely.

Conclusion

The current study has provided initial evidence for the reliability, validity and clinical feasibility of Chinese version of the Doloplus-2 scale among cognitively impaired older people with communication difficulty; subsequently, further work is needed to replicate these findings. Furthermore, based on the study results, we feel confident in encouraging institutional caregivers to integrate the scale into routine clinical practice in the long-term care facilities, but the clinical effect of the scale for effectively managing pain should be further examined.

References

AGS panel on persistent pain in older persons (2002) The management of persistent pain in older persons. *American Geriatrics Society* **50**, s205-s224.

Alexopoulos G..S., Abrams R.C., Young R.C. & Shamoian C.A. (1988) Cornell scale for depression in dementia. *Society of Biological Psychiatry* **23**, 271-284.

Bjoro K. & Herr K. (2008) Assessment of pain in the nonverbal or cognitively impaired older adult. *Clinics in Geriatric Medicine* **24**, 237-262.

Blomqvist K. & Hallberg I. (2001) Recognising pain in older adults living in sheltered accommodation: the views of nurses and older adults. *International Journal of Nursing Studies* **38**, 305-318.

Cha E., Kim K. & Erlen J. (2007) Translation of scales in cross-cultural research: issues and techniques. The Journal of Advanced Nursing **58**. 386-395.

Cohen-Mansfield J. (1986) Agitated behaviors in the elderly: II. Preliminary results in the cognitively deteriorated. *Journal of American Geriatrics Society* **34**, 722-727.

Cohen-Mansfield J., Marx M.S. & Rosenthal A.S. (1989) A description of a agitation in a nursing home. *Journal of Gerontology* **44**, M77-84.

Field A. (2005) *Discovering Statistics Using SPSS.* SAGE publications, London.

Firris L. & Norton P. (1992) Basic concepts in reliability and validity. In *Tools for Primary Care Research* (Stewart M. Tudiver F. Bass M. Dunn E. & Norton P., eds), Sage, Newbury Park, CA, USA, pp.64-76.

Folstein M., Folstein S.E. & McHugh P. (1975) Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**, 189-198.

Green J. & Young J. (2001) A test-retest reliability study of the Barthel Index, the Rivermead mobility Index, the Nottingham extended activities of daily living scale and the Frenchay activities index in stroke patients. *Disability and Rehabilitation* **23**, 670-676.

Hadjistavropoulos T., Voyer P., Sharpe D., Verreault R. & Aubin M. (2008) Assessing pain in demenia patients with comorbid delirium and/or depression. *Pain Management Nursing* **9**, 48-54.

Herr K., Bjoro K. & Decker S. (2006) Tools for assessment of pain in nonverbal older adults with dementia : a state-of the science review. *Journal of Pain and Symptom Management* **31**, 170-192.

Holen C., Saltvedt I., Fayers M., Bjornnes M., Stenseth G., Hval B., Fibet M., Loge J. & Kaasa, S. (2005) The Norwegian Doloplus-2, a tool for behavioural pain assessment : translation and pilot – validation in nursing home patients with cognitive impairment. *Palliative Medicine* **19**, 411-417.

Holen J., Saltvedt I., Fayers P., Hjermstad M., Loge J. & Kaasa S. (2007) Doloplus-2, a valid tool for behavioural pain assessment? *BMC Geriatrics* **7**, 1-9.

Hsueh I.P., Lee M.M. & Hsieh C.L. (2001) Psychometric characteristics of the Barthel activities of daily living index in stroke patients. *Journal of the Formosan Medical Association* **100**, 526-531.

Hyrkas K., Appelqvist-Schmidlechner K. & Oksa L. (2003) Validating an instrument for clinical supervision using an expert panel. International Journal of Nursing Studies **40**, 619-625.

Juva K., Sulkava R., Erkinjuntti T., Ylikoski R., Valvanne J. & Tilvis R. (1994) Staging the severity of dementia: comparison of clinical (CDR, DSM-III-R), functional (ADL, IADL) and cognitive (MMSE) scales. *Acta neurologica Scandinavica* **90**, 293-298.

Lafebvre-Chapiro L. & Doloplus group (2001) The DOLOPLUS-2 scale: evaluating pain in the elderly. *European Journal of Palliative care* **8**, 191-194.

-194.

Lam C.K., Lim P.P.J., Low B.L., Ng L., Chiam P.C. & Sahadevan S. (2004) Depression in dementia: a comparative and validation study of four brief scales in the elderly Chinese. *International Journal of Geriatric Psychiatry* **19**, 422-428.

428.

Lin L.C., Kao C.C., Tzeng Y.L., Lin Y.J. (2007) Equivalence of Chinese version of the Cohen-Mansfield Agitation Inventory. *Journal of Advanced Nursing* **59**, 178-185.

Mahoney F.I. & Barthel D.W. (1965) Functional evaluation: the Barthel index. *Maryland State Medical Journal* **14**, 61-65.

Maneesriwongul W. & Dixon J. (2004) Instrument translation process: a method review. *Journal of Advanced Nursing* **48**, 175-186.

Miller R.J., Snowdom J. & Vaughan R. (1995) The use of the Cohen- Mansfield agitation inventory in the assessment of behavioral disorders in nursing homes. *Journal of American Geriatrics Society* **43**, 546-549.

LoBindo-Wood G.. & Haber J. (1998) *Nursing Research: Methods, Critical Appraisal, and Utilization*, 3rd edn. Mosby, St Louis, MO, USA, pp.327-350.

Louis, MO, USA, pp.327-350.

, pp.327-350.

Pautex S., Herrmann F., Lous L., Fabjan M., Michel J. & Gold G. (2005) Feasibility and reliability of four pain self-assessment scales and correlation with an observational rating scale in hospitalized elderly demented patients. *The Journals of Gerontolog* **60**, 524-529.

Pautex S., Herrmann F., Michon A., Giannakopoulos P., Michel J. & Gold G. (2007) Psychometric properties of the Doloplus-2 observational pain assessment scale and comparison to self-assessment in hospitalized elderly. *The Clinical Journal of Pain* **23**, 774-778.

8.

Sainsbury A., Seebass G., Bansal A. & Young J.B. (2005) Reliability of the Barthel Index when used

with older people. *Age and Ageing* **34**, 228-232.

Smith M. (2005) Pain assessment in nonverbal older adults with advanced dementia. *Perspectives in Psychiatric Care* **41** , 99-113.

Snow A.L., O'Malley K.J., Cody M., Kunik M.E., Ashton C.M., Beck C., Bruera E. & Novy D. (2004) A conceptual model of pain assessment for noncommunicative persons with dementia. *The Gerontologist* **44**, 807-817.

Stolee P., Hillier L., Esbaugh J., Bol N., Mckellar L. & Gauthier N. (2005). Instruments for the assessment of pain in older persons with cognitive impairment. *American Geriatrics Society* **53**, 319-326.

Tombaugh T.N., McIntyre N.J. (1992) The Mini-Mental State Examination: a comprehensive review. *Journal of the American Geriatrics Society* **40**, 922–35.

Weiner D., Peterson B. & Keefe F. (1999) Chronic pain-associated behaviors in the nursing home: resident versus caregiver perceptions. *Pain* 80, 577-588.

Yip P.K., Shyu Y.I., Liu S.I., Lee J.Y., Chou C.F. & Chen R.C. (1992) An epidemiological survey oh dementia among elderly in an urban district of Taipei. *Acta Neuroglogica Sinica* **1**, 347-354.

-354.

4.

Zwakhalen S., Hamers J., Abu-Saad H. & Berger M. (2006) Pain in elderly people with severe dementia : a systematic review of behavioural pain assessment tools. *BioMed Central Geriatrics* **6**, 1-15.

ZwakhalenG., Hamers H. & Berger F. (2006) The psychometric quality and clinical usefulness of three pain assessment tools for elderly people with dementia. *Pain* **126**, 210-220.

Table 1. Demographics of elders with moderate and severe dementia and RNs[a]

characteristics	*N* (%)	Mean ± SD	characteristics	*N* (%)	Mean ± SD
Older people with morderate and severe dementia (n=*241*)					
Age (yrs)		79.28±9.42	Numbers of falls during the recent six months		0.50±1.75
Gender			Folstein MMSE [b]		5.26±5.46
Female	118(49)		MMSE≦ 10 (%)	185(76.8)	
Male	123(51)		Disability		34±31.37
Religion			Cornell scale		1.59±1.61
Buddhism/Taoism	121(50)		CMAI [c]		35.34±6.83
Christianity/Catholicism	37(16)				
Others	10(4)		*RNs* (n=14)		
Non	73(30)		Age		35.29±8.17
Education			Religion		
Uneducated	85(35)		Buddhism/Taoism	7(50)	
Primary school	60(25)		Christianity/Catholicism	3(21)	
Secondary school	35(15)		Non	4(29)	
Advanced level	29(12)		Education		
University	29(12)		Vocational nursing school	1(7)	
Graduate degree	3(1)		Advanced level	10(71)	
Length of stay in institution		38.32±32.37	University	3(21)	
Medical conditions			Working experience (yrs)		16.83±19.10
Number of diagnoses		2.23±1.51	Working in the institution (yrs)		4.09±3.22
Number of medications		6.47±3.49	Average working time per week		44±3.92
Number of analgesics		0.25±0.55			
Past pain history					
Experiencing pain	164(68)				
No pain	77(32)				

[a]Registered nurses

[b]Mini-Mental State Examination

[c]Cohen-Mansfield Agitation Inventory

Table 2. Distribution and reliability of the C-Doloplus-2 scores (N=241)

	RNs[a]				
Item	Mean± SD	Range	Cronbach's α	Alpha if item is deleted	inter-rater reliability (ICC[b])
Somatic reactions	**1.25±2.12**	0-13	0.79		0.79
Verbal complaints	0.29±0.64			0.73	
Protective body posture	0.29±0.65			0.71	
Self-protection of sore ar	0.29±0.66			0.70	
Facial expression	0.29±0.67			0.72	
Sleep pattern	0.29±0.68			0.72	
Psychomotor reactions	**2.1±1.86**	0-6	0.87		0.84
Washing and/or dressing	1.16±1.01			0.72	
Mobility	0.94±0.97			0.73	
Psychosocial reactions	**1.12±1.88**	0-8	0.67		0.60
Communication	0.31±0.77			0.71	
Social life	0.62±0.97			0.72	
Behavioral problems	0.20±0.65			0.72	
Total C-Doloplus-2	**4.49±4.10**	0-21	0.74		0.81

[a]Registered Nurse

[b]Intra-class correlations.

Table 3. Pearson correlations between C-Doloplus-2 and known correlates of pain

Variable	C-Doloplus-2		
	MMSE≦ 17	11≦ MMSE≦ 17	MMSE≦ 10 (*N*=185)
Past pain history	0.08	0.05	0.08
Numbers of falls during	0.02	0.06	0.02
Disability	-0.28*	-0.38**	-0.22**
Cornell scale	0.12	0.08	0.12*
CMAI[a]	0.04	-0.07	-0.07

[a]Cohen-Mansfield Agitation Inventory

*P<0.05 **p<0.01

are not consistent in presenting means and standard deviations – what is the convention of the journal you intend to submit to?

Re-state hypothesis

Do not understand what you mean here