Assessing depression in patients with chronic pain: A comparison of three rating scales

Running title: Assessing depression in chronic pain

Wong, W. S.,^a Chen, P. P.,^b Yap, J.,^b Mak, K. H.,^c Tam, B. K. H.,^a & Fielding, R.^d

^a Department of Psychological Studies, The Hong Kong Institute of Education, 10 Lo Ping Road, Tai Po, Hong Kong.

^b Department of Anaesthesiology & Operating Services, Alice Ho Miu Ling Nethersole Hospital, Tai Po, Hong Kong.

^c Department of Orthopedics & Traumatology, Kwong Wah Hospital, Kowloon, Hong Kong.

^d Health Behaviour Research Group, School of Public Health, The University of Hong Kong, Pokfulam, Hong Kong.

Abstract

Background: Considerable evidence has suggested depression is significantly more prevalent in patients with chronic pain. A number of studies exclusively based on Western samples have evaluated the effectiveness of depression rating scales in assessing depression in the chronic pain context. The objective of this cross-sectional study was to compare within a Chinese chronic pain sample three depression rating scales commonly used in identifying depression.

Methods: A total of 366 Chinese patients with chronic pain attending an orthopedics specialist clinic (n=185) and a multidisciplinary pain clinic (n=181) in Hong Kong completed a structured interview using CIS-R and two depression rating scales, the Beck Depression Inventory (BDI standard and short form) and the Center for Epidemiological Studies --- Depression (CES-D). Patient scores on the BDI and CES-D were then assessed against their responses on the CIS-R to determine their effectiveness.

Results: The prevalence of depression was 20.2% and 57.8% in the Orthopedics and Pain Clinic sample respectively. Results of ROC analyses showed that all the three measures performed well at predicting depression with AUC ≥ 0.89 and high sensitivity and specificity.

Conclusions: Our findings suggest that the three depression measures assessed have good predictive validity in the Chinese chronic pain context, and they could be used as screening or diagnostic measures of depression in Chinese chronic pain patients. The decision of using a specific measure and a specific cutoff score should be based on study aim and setting.

Keywords: Depression; Chronic pain; BDI; CES-D; Chinese.

1. Introduction

The prevalence of depression among patients with chronic pain is significantly higher at 30-54% than in the general population (5%) (APA, 2006; Blazer et al., 1994; Demyttenaere et al., 2007), whereas about 43.4% (Ohayon, 2004)to 100% (Ward et al., 1979) of depressed patients report pain. Depression can hamper the management of chronic pain. Depression predicted poor responses to rehabilitation program (Harter et al, 2002), and higher dose and administration of opioids (Jensen et al., 2006) among chronic pain patients. Pain has been found to predict subsequent depression (Breslau et al., 1994). However, current or previous psychiatric disorder also predicted pain (Silberstein, 2001). The depression-pain link appears to be stronger in older patients with chronic pain (Carroll et al., 2000; Patten et al., 2006; Turk et al., 1995). Depression was also associated with higher rates of health care utilization among patients with chronic pain (Tripp et al., 2006).

Despite the quality of life and economic impacts, the detection and treatment of depressed people in the chronic pain population is understudied. The US Preventive Services Task Force (USPSTF) reviewed randomized controlled trials that examined the efficacy of screening for depression in primary care settings and found good evidence that screening decreased clinical morbidity as well as increased detection by a factor of two to three (Preventive Services Task Force, 2002a). Consequently, the USPSTF recommends routine screening for depression of all adults in primary care, arguing that the benefits of screening are likely to outweigh any potential harm. In a population at high risk of depression, such as chronic pain patients, screening for depression therefore should be even more beneficial.

Several tools have been assessed for their diagnostic efficiency to detect depression or depressive symptoms based on DSM criteria: (1) the Beck Depression Inventory-Standard and Short Forms (BDI/BDI-SF) (Geisser et al., 1997; Love, 1987; Turner and Romano, 1984), the Centre for Epidemiological Studies-Depression scale (CES-D) (Geisser et al., 1997), the MMPI Depression scale (Love, 1987; Turner and Romano, 1984), the Zung Self-Rating Depression Scale (ZSRDS) (Turner and Romano, 1984), and the Middlesex Hospital Questionnaire Depression scale (MHQ-D) (Love, 1987). The MMPI Depression scale seems to lack accuracy in patients with chronic pain (Turner and Romano, 1984) and chronic low back pain (Love, 1987). Similar low diagnostic efficiency was also reported for the MHQ-D (Love, 1987). The optimal cutoff score of 49/50 was identified for ZSRDS (sensitivity 83%, specificity 81%) (Turner and

Romano, 1984). The BDI gave cutoffs of 11/12 in 68 chronic low back pain patients (Love, 1987), 12/13 in 40 chronic pain patients (Turner and Romano, 1984), and 20/21 in 132 consecutive chronic pain patients attending pain management programs (Geisser et al., 1997), giving optimal cutoff for the BDI-SF as 7/8 (Turner and Romano, 1984). As for the CES-D, optimal cutoff is proposed to be 26/27 (Geisser et al., 1997). Cutoff score discrepancies are partly due to sample and methodological heterogeneity.

The BDI and CES-D have been validated in Chinese with good validity and reliability (Shek, 1991; Ying, 1988) but their efficiency and accuracy in assessing depression in the Chinese chronic pain population remains unknown. Because the existing cutoffs were obtained in Western populations, their applicability in the Chinese context is unknown. In this cross-sectional study, we compared the efficiency of the BDI, BDI-SF, and the CES-D as screening tools for depression in a sample of Hong Kong-Chinese patients with chronic pain. Specialist services for chronic pain in Hong Kong are currently limited to eight multidisciplinary pain clinics in public hospitals. Most patients with chronic pain are managed by orthopedics specialists. A recent study shows that the rates of depression in the Pain Clinic (57.1%) were significantly higher than those in the Orthopedics sample (20.2%) (Wong and Fielding, accepted). In light of this, chronic pain patients attending both types of specialist services were assessed in this study.

2. Methods

2.1. Study sample

After approval from both the university and hospital IRBs, consecutive patients attending for musculoskeletal pain problems at an orthopedic specialist out-patient clinic and a pain specialist clinic of two Hong Kong public hospitals were invited to participate in the present study. Patients who were (1) \geq 18 years of age; (2) native Cantonese speakers; (3) lacking communication problems or physical conditions preventing completion of the study measures; (4) exhibiting no confusion and having no prior diagnosis of cognitive impairment from medical records; and (5) willing to participate in the study and to give written consent.

A total 370 patients from the Orthopaedics outpatient clinics (n=185) and the Pain Clinic (n=185) participated. Four patients from the Pain Clinic did not complete the depression measures and were therefore excluded from analyses (Table 1). In both samples, women constituted over half of the sample in both clinics (\geq 51.9%). Pain Clinic participants were older

(mean=43.02 vs 39.05 years, *SD*=11.11; *t*=-3.53, *p*<0.001) and more (61.8%) reported low monthly household income (<HK\$15,000; χ^2 =10.37, *p*<0.05) compared to the Orthopaedics participants. More Orthopaedic clinic participants (39.5% vs. 24.3%) were never married (χ^2 =12.22, *p*<0.01), or had attained tertiary education (12.0% vs. 6.1%; χ^2 =11.64, *p*<0.05). While more Orthopaedic clinic participants were in full-time employment (59.6% vs. 36.7%), more than twice as many Pain Clinic participants reported that they were unemployed (39.2% vs 16.4%) (χ^2 =40.08, *p*<0.001).

2.2. Procedures

Patients were recruited during visits for clinical consultations with doctors. In each targeted clinic session, where manpower permitted, every attending patient was invited to participate. Individual face-to-face interviews were conducted before or after clinical consultation among patients meeting chronic pain criteria (pain duration >3 months) and fulfilling the inclusion criteria. The interviews were carried out by trained interviewers who had at least 1 year prior interviewing experience and had undertaken a 1-day training programme in the use of the study measures. Fieldwork was closely monitored by the principal investigator of the project.

2.3. Measures

Depression diagnostic interview

The Revised Clinical Interview Schedule (CIS-R), a structured diagnostic interview devised for use by non-medical professionals, was employed to assess depression morbidity for all participants in this study (Lewis et al., 1992). The CIS-R consists of 14 sections, assessing 14 neurotic symptoms in the week before the interview: somatic symptoms, fatigue, concentration and forgetfulness, sleep problems, irritability, worry about physical health, depression, depressive ideas, worry, anxiety, phobias, panic, compulsion and obsession. Each section is scored on a 0-4 scale (except depressive ideas 0-5). The rating obtained at interview can be presented for each symptom group, and they can be summed to yield an overall score, which is taken to indicate the severity of any minor psychiatric disorder. Current depressive episodes were diagnosed from CIS-R scores by applying algorithms of the ICD-10 system classification (Singleton et al., 2001). Three types of depressive episode diagnoses, including mild, moderate, and severe, were grouped as Depressive Disorders category.

Beck Depression Inventory (BDI)

The full 21-item version of the Beck Depression Inventory (BDI) was designed to assess cognitive, behavioral and somatic symptoms of depression (Beck et al., 1961), generating total scores ranging from 0 to 63. Scoring on a 0-3 Likert scale, BDI has demonstrated good internal consistency in both psychiatric (mean Cronbach's $\alpha = 0.86$) and non-psychiatric patients (mean Cronbach's $\alpha = 0.81$) (Beck et al., 1988). The Chinese version of the BDI (C-BDI) also showed good convergent validity with psychological well-being (Shek, 1991). Previous study indicated that a cut-off of 20/21 yielded adequate specificity (68.2%) in outpatients attending a pain management programme (Geisser et al., 1997). The use of a cut-off of 13/14 has been suggested for assessing depression among Chinese patients with chronic pain (Lee et al., 2008).

Beck Depression Inventory Short Form (BDI-SF)

To minimize scale contamination in patients with somatic symptoms, the 13-item BDI Short Form (BDI-SF) was derived from the original 21-item scale by excluding 8 somatic items (Beck and Beck, 1972). This short version possessed good internal consistency (Cronbach's α = 0.83) and demonstrated high correlation with the original standard version (*r*s ranging 0.93-0.96) (Beck and Beck, 1972; Reynolds and Gould, 1981). A study utilizing both in- and out-patients with chronic pain suggested a cut-off of 7/8 for the use of the BDI-SF (Turner and Romano, 1984).

Center for Epidemiological Studies-Depression Scale (CES-D)

The 20-item Center for Epidemiological Studies-Depression Scale (CES-D) was developed for screening community depressive symptomatology (Radloff, 1977). Rating on a 4-point scale, CES-D demonstrated good internal consistency (Cronbach's α ranged 0.84-0.90). The Chinese version of the CES-D has good psychometric properties (Cronbach's α =0.77, a four-factor model explained 48% of the total variance) (Ying, 1988). With a cutoff of 26/27, the CES-D demonstrated adequate sensitivity (81.8%) in a sample of outpatients with chronic pain (Geisser et al., 1997).

Chronic Pain Grade (CPG)

The 7-item Chronic Pain Grade questionnaire (CPG) was employed in this study to evaluate pain characteristics of the present sample. The CPG assesses three domains of pain severity: persistence, intensity and disability/interference.(Von Korff et al., 1992) Three intensity items ask respondents to rate their current, average and worst pain intensity on 0-10 Numerical Rating Scales (NRS) (0="No pain at all"; 10="Pain as bad as could be"). A Characteristic Pain Intensity Score is derived by averaging the responses to the intensity items and multiplying this by 10. Three CPG items assess pain interference with (1) daily activities, (2) social activities, and (3) working ability using 0-10 NRSs. The CPG Disability Score is derived by multiplying the average of the three interference items by 10. Persistence is assessed in the original CPG by asking the respondent to indicate the number of days out of the past three months days that he/she was disabled by pain. The Disability Score and the number of disability days are recoded into 5point scales and summed, yielding "Disability Points". Based on the Pain Intensity Score and Disability Points, the CPG classifies respondents into five hierarchical grades (refer to Table 1 for the five hierarchical grades). The English version of the CPG possesses good psychometric properties (Smith et al., 1997) and is responsive to change in pain severity over time (Elliott et al., 2000). The underlying structure of the CPG among Chinese clusters into 3 main dimensions: Disability, Intensity, and Persistence (Fielding and Wong, 2008). Cronbach's as for the CPG Disability and Characteristic Intensity scales were .87 and .68, respectively.

To control for potential order effects, three questionnaires forms were created and administered randomly: Form A (CIS-R, BDI/BDI-SF, and CES-D), Form B (BDI/BDI-SF, CIS-R, and CES-D), and Form C (CES-D, BDI/BDI-SF, and CIS-R). Comparisons of patients administered to the three questionnaire forms showed no statistically significant differences on the proportion of depressed cases as identified by the CIS-R and on the mean scores of BDI, BDI-SF, and CES-D (all p>0.05).

2.4. Data analysis

The sample was categorized into cases or noncases of depression based on the CIS-R diagnosis; thus, any participant with an ICD-10 diagnosis of depressive episodes was identified as a "case". To evaluate the BDI, BDI-SF and CES-D performance we determined the sensitivity (the percentage of true "cases" identified by the instrument), specificity (the percentage of true "noncases" identified by the instrument), positive predictive value (PPV; the proportion of all

those tested as positive who were correctly identified as such), negative predictive value (NPV; the proportion of all those tested as negative who were correctly identified as such), and overall efficiency (Eff; the proportion of correct classifications) using the CIS-R diagnosis as the gold standard criterion. The receiver operating characteristic (ROC) curve, a plot of sensitivity against 1-specificity at different cutoffs, helped to identify the optimal cutoff score, usually when an increase in sensitivity is associated with a sharp drop in specificity. The optimal cutoff score was determined where the sum of sensitivity and specificity is maximized. The area under the curve (AUC) was also calculated, with the higher AUC indicating better balance between sensitivity and specificity. Prior to the ROC analyses, a series of univariate regression analysis was run to evaluate the association of sociodemographic factors with prevalence, pain outcomes, and depression. The results yielded no significant association between any of the 7 sociodemographic factors were not controlled for in the ROC equations. Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 15.0 (SPSS Inc, 2002).

3. Results

3.1. Pain characteristics of the sample

Table 1 reports the sample pain characteristics. The two samples differed markedly in their clinical characteristics. Significantly almost twice as many Pain Clinic participants reported pursuing pain-related litigation (χ^2 =14.76, p<0.001) and medico-legal compensation (χ^2 =17.64, p<0.001) than did their Orthopedics counterparts. The proportion of patients that cited pain as the main reason for their first clinic visit was 91.8% for the Orthopedics Clinic and 82.5% for the Pain Clinic. Moreover, compared to just 6% of Orthopedic Clinic participants, more Pain Clinic participants (15.3%) indicated pain was one, but not the main symptom driving their first clinic visit (χ^2 =8.28, p<0.05). Pain Clinic participants reported a duration of chronic pain averaging 5.31 years/1938 days (*SD*=6.73 years/2455 days; median=3 years/1,095 days), over 50% longer than the average duration reported by Orthopaedics participants (mean=2.55 years/929 days; *SD*=3.61 years/1316 days; median=1 year/365 days) (*t*=-4.87, p<0.001). Most Orthopaedics participants (59.6%) had suffered from chronic pain for seq years, while 10.4% of the Pain Clinic participants reported having had chronic pain for more than 10 years. The number of pain sites reported by Pain Clinic participants was significantly higher than that reported by Orthopaedics

participants (t=7.88, p<0.001). While over half of the Orthopaedics Clinic participants had only one pain site, 11.7% of the Pain Clinic participants had 6 or more pain sites.

Pain Clinic participants reported greater pain intensity than their Orthopaedics counterparts (all p<0.001). The two samples also significantly differed on pain interference measures, with Pain Clinic participants scoring significantly poorer on daily activities (t=-6.91, p<0.001), social activities (t=-7.76, p<0.001), and working abilities (t=-5.35, p<0.001). Compared to Orthopaedics Clinic patient, Pain Clinic participants reported a significantly greater mean number of pain-associated disability days (t=-6.82, p<0.001) and pain-associated leave of absence (t=-3.40, p<0.01). The CPG classified 57.9% of Orthopaedics participants as Grade II or below (high pain intensity but low related disability), while 32.8% and 47.2% of Pain Clinic participants were classified as Grade III and IV respectively. Significantly more Pain Clinic than Orthopaedics participants achieved a higher CPG classification (χ^2 =67.18, p<0.001).

Based on ICD-10 diagnostic criteria, significantly more patients in the Pain Clinic (57.8%) were identified as having depressive disorder as compared to the Orthopaedics sample (χ^2 =53.89, p<0.001). Significant differences on the three depression measures were also found, with the Pain Clinic sample scoring significantly higher on all three measures than their Orthopaedics counterparts (all p<0.001). As shown in Table 2, depressed patients in both clinics and the entire sample scored significantly higher all three depression measures (all p<0.001).

3.2. The performance of BDI in assessing depression

The screening test indices for the Orthopaedics Clinic, the Pain Clinic, and the entire sample are listed in Table 3. The corresponding ROC curves are shown in Figure 1 to Figure 3 respectively. The ROC AUC for the BDI for the Orthopaedics, the Pain Clinic, and the entire sample were 0.93 (95% CI, 0.89-0.98), 0.89 (95% CI, 0.85-0.94), and 0.92 (95% CI, 0.87-0.98) respectively. The optimal cutoff score for the Orthopaedics sample was 12/13 yielding a sensitivity of 84%, specificity of 87%, PPV of 62%, NPV of 95%, and Eff of 86%. As for the Pain Clinic sample, a cutoff score of 15/16 was the optimal score producing a sensitivity of 84%, specificity of 72%, PV of 80%, NPV of 76%, and Eff of 79%. The recommended cutoff score for the combined sample was also 15/16, which produced 81% sensitivity, 86% specificity, 78% PPV, 88% NPV, and 84% Eff.

3.3. The performance of BDI-SF in assessing depression

The ROC AUC for the BDI-SF for the Orthopaedics, the Pain Clinic, and the entire sample were 0.92 (95% CI, 0.86-0.97), 0.89 (95% CI, 0.85-0.94), and 0.92 (95% CI, 0.90-0.95), respectively. A cutoff score of 7/8 was recommended for the Orthopaedics sample. This cutoff obtained a sensitivity of 86%, specificity of 85%, PPV of 60%, NPV of 96%, and Eff of 86%. The optimal cutoff score of 9/10, which produced 85% sensitivity, 70% specificity, 80% PPV, 78% NPV, and 79% Eff, was suggested for the Pain Clinic sample. For the entire sample, the optimal cutoff of 8/9 yielded a sensitivity of 88%, specificity of 81%, PPV of 75%, NPV of 91%, and Eff of 84%.

3.4. The performance of CES-D in assessing depression

The ROC AUC for the CES-D for the Orthopaedics, the Pain Clinic, and the entire sample were 0.92 (95% CI, 0.87-0.98), 0.91 (95% CI, 0.86-0.95), and 0.93 (95% CI, 0.91-0.96) respectively. For the Orthopaedics sample, the recommended cutoff score was 21/22, producing 84% sensitivity, 91% specificity, 70% PPV, 96% NPV, and 90% Eff. The optimal cutoff score of 26/27 was obtained for the Pain Clinic sample. This cutoff obtained a sensitivity of 85%, specificity of 79%, PPV of 85%, NPV of 80% and Eff of 83%. As for the entire sample, a cutoff score of 26/27 yielded a sensitivity of 81%, specificity of 91%, PPV of 86%, NPV of 88%, and Eff of 87%.

4. Discussion

We believe this is the first study to evaluate the relative screening efficiency of the BDI, BDI-SF, and CES-D, in chronic pain among Chinese patients. The prevalence of depression based on the ICD-10 criteria was 20.2% and 57.8% for the Orthopaedics and Pain Clinic sample respectively. The AUCs obtained for the three studied measures were high (≥ 0.92) for the Orthopaedics sample and moderate-to-high (≥ 0.89) for the Pain Clinic sample.

The ROC analyses indicated that all three measures demonstrated sufficiently good sensitivity and specificity in detecting depression in Chinese chronic pain patients attending orthopaedics or multidisciplinary pain services. Optimal cutoff scores for BDI, BDI-SF, and CES-D for the Orthopaedics sample were 12/13, 7/8, and 21/22 respectively. For the Pain Clinic sample, the optimal cutoff scores were 15/16, 9/10, and 26/27 for BDI, BDI-SF, and CES-D

respectively. Higher cutoffs in the Pain Clinic sample indicate a higher threshold is needed to detect depression among attendees of multidisciplinary pain services. This is possibly because patients attending multidisciplinary pain services had more severe pain problems for longer, and/or coped more effectively with their pain, and because the chronic pain population is heterogenous, even within the same cultural group. Despite differences in optimal cutoff scores and variations in sensitivity and specificity between the two samples, our results are consistent with the findings reported in a study that most of the screening instruments for depression had good sensitivity but only fair specificity (Preventive Services Task Force, 2002b).

The cutoff scores obtained in this Chinese sample are somewhat different from those reported previously in Western samples. The optimal BDI-SF cutoff (7/8) in the present Orthopaedics sample is the same as Turner and Romano's sample (Turner and Romano, 1984), though a higher cutoff was found in the present Pain Clinic sample (9/10). Considering the CES-D, the optimal cutoff score of 26/27 was obtained in our Pain Clinic sample, which is the same as Gessier et al.'s (Geisser et al., 1997) recommendation. However, the optimal CES-D cutoff found in our Orthopaedics sample was lower (21/22) than that of Gessier et al. (Geisser et al., 1997). As for the BDI, the optimal cutoff score in our Orthopaedics sample (12/13) was the same as that recommended by Turner and Romano (Turner and Romano, 1984) and close to that suggested by Love (11/12) (Love, 1987). Yet, the BDI cutoff score reported by Gessier et al. (20/21; 68% sensitivity; 78% specificity) (Geisser et al., 1997) was much higher than both of our Orthopaedics (12/13) and Pain Clinic (15/16) sample.

One possible explanation for the differences on optimal cutoff scores between this study and the previous studies may be due to differences between the samples on the mean scores of the rating measures. While no significant differences were found between our samples and the sample of Turner and Romano's study (Turner and Romano, 1984) on BDI and BDI-SF scores, results of our post-hoc analyses showed that the depressed patients in Gessier et al's (Geisser et al., 1997) study scored significantly higher on BDI (mean=25.50) but lower on the CES-D than our Orthopaedics sample (BDI: mean=20.86, t=2.25, p<0.05; CES-D: mean=32.80, t=3.13, p<0.01). Nondepressed patients in Gessier et al.'s (Geisser et al., 1997) study had significantly higher BDI (mean=14.80) and CES-D (mean=19.40,) scores than our nondepressed Orthopaedics (BDI: mean=6.50; t=9.36, p<0.001; CESD: mean=10.27; t=6.79, p<0.001) and nondepressed Pain Clinic (BDI: mean=10.56; t=3.44, p<0.001) sample. While these observed mean differences may suggest genuine cultural differences between the present Chinese sample and the two samples in the United States (Geisser et al., 1997; Turner and Romano, 1984), they may also be arisen from cross-cultural measurement artefacts such as response set biases. For instance, research has consistently shown that Chinese and other Asians tend to avoid extremes and prefer the midpoints of scales, a response style which is influenced by cultural virtues of promoting moderation and non-judgementalism (Chen et al., 1995; Zax and Takahashi, 1967). Future cross-cultural and/or international comparison on the cutoff and means scores of the three depression measures assessed in this study should take these issues into consideration.

Furthermore, sample heterogeneity across studies may also contribute to different optimal cutoff scores obtained. Our Pain Clinic sample reported significantly longer pain duration (mean=64.62) than that of Gessier et al's (Geisser et al., 1997) sample (mean=44.20; t=2.52, p<0.05) whilst the nondepressed sample in Turner and Romano's (Turner and Romano, 1984) study (mean=51.61) was significantly younger than both of our samples (Orthopaedics: mean=38.10; *t*=5.24, *p*<0.001; Pain Clinic: mean=42.84, ;*t*=3.22, *p*<0.01). The proportion of female patients in Gessier et al's (60%) (Geisser et al., 1997) study was also significantly higher than both of our samples (52%; χ^2 =11.98, *p*<0.001; Pain Clinic: 55%, χ^2 =8.26, *p*<0.01).

The recommended cutoff scores in this report were based on an optimal balance between sensitivity and specificity. We do not suggest strict adherence to one specific cutoff score for the three depression measures assessed. The efficiency of any test depends on the sensitivity and specificity derived from the cutoff scores employed to define a condition. As sensitivity increases, specificity tends to decline; conversely, a more specific test tends to be less sensitive (Mulrow et al., 1995; Whooley et al., 1997). Furthermore, the purpose and scope of applying a tool affects the adoption of a cutoff score. If evaluating an instrument as a diagnostic tool, a high sensitivity and a high NPV would offer important information facilitating the efficiency of conducting initial diagnostic evaluation. If evaluating an instrument as a screening tool, a high specificity and a high PPV should be sought. As the PPV values depend on the prevalence of disease, if a screening tool is applied in disease with relatively low prevalence, the corresponding PPV would have a lower value, and vice versa. The high prevalence of depression in chronic pain populations makes these instruments highly appropriate for screening and the amenability of depression to treatment makes screening for depression highly desirable and effective.

Compared to previous studies in chronic pain samples (Geisser et al., 1997; Love, 1987; Turner and Romano, 1984), the bigger sample size in this study is a strength, conferring as it does adequate statistical power to the analyses. The examination of the three rating scales' performances in detecting depression in patients with chronic pain attending different types of pain services addressed the question of possible patient heterogeneity in scale performance, an issue not previously examined. The interpretation of our data should take into account three methodologicallimitations. First, CIS-R was employed as a gold standard in this study. The CIS-R has shown to have a sensitivity of 100% and specificity of 88% for diagnosis of depressive episode (Jordanova et al., 2004) when correctly administered. The CIS-R demonstrated good interrater reliability of the CIS-R (median kappa=0.88), and pairs of psychiatrists were shown to be only slightly more reliable (kappa=0.75) than nonprofessional psychiatric pairs (kappa=0.70) in determining "cases" using the CIS-R (Lewis et al., 1992). These findings support the use of the CIS-R as a gold standard for assessing depression. The measure and its variants have also been extensively used in different countries and regions worldwide (Blazer et al., 1994; Carter et al., 2001; Henderson et al., 2000; Jenkins et al., 1997; WHO, 2000). Second, interrater reliability for the CIS-R interview was not determined in this study, which may hamper the results obtained on PPV. Finally, the current study assessed only three depression scales. Other depression scales, such as the Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith, 1983) which is widely used in non-psychiatric population, is not examined in this study, leaving their diagnostic efficiency in detecting depressive symptoms among patients with chronic pain unknown. This should be addressed in future studies.

Despite these limitations, our data offer preliminary support for the utility of the BDI, BDI-SF, and CES-D for screening for depression among Chinese patients with chronic pain. Factors such as study aim and setting (e.g., orthopaedics vs multidisciplinary pain services) should be taken into consideration when deciding which specific measure and/or specific cutoff score to be used.

Reference

- APA, A.P.A., 2006. Diagnostic and Statistical manual of Mental Disorder. 3rd Edition, Revised. American Psychiatric Association, Washington, DC
- Beck, A.T., Beck, R.W., 1972. Screening depressed patients in family practice. A rapid technic. Postgraduate Medicine 52, 81-85.
- Beck, A.T., Steer, R.A., Garbin, M.G., 1988. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review 8, 77-100.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. Archives of general psychiatry 4, 561-571.
- Blazer, D.G., Kessler, R.C., McGonagle, K.A., Swartz, M.S., 1994. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. Am J Psychiatry 151, 979-986.
- Breslau, N., Davis, G.C., Schultz, L.R., Peterson, E.L., 1994. Joint 1994 Wolff Award Presentation. Migraine and major depression: a longitudinal study. Headache 34, 387-393.
- Carroll, L., Cassidy, J.D., Cote, P., 2000. The Saskatchewan health and back pain survey: The prevalence and factors associated with depressive symptomatology in Saskatchewan adults. Canadian Journal of Public Health 91, 459-464.
- Carter, R.M., Wittchen, H.U., Pfister, H., Kessler, R.C., 2001. One-year prevalence of subthreshold and threshold DSM-IV generalized anxiety disorder in a nationally representative sample. Depression and Anxiety 13, 78-88.
- Chen, C.S., Lee, S.Y., Stevenson, H.W., 1995. Response style and cross-cultural comparisons of rating scales among East Asian and North American Students. Psychological Science 6, 170-175.
- Demyttenaere, K., Bruffaerts, R., Lee, S., Posada-Villa, J., Kovess, V., Angermeyer, M.C., Levinson, D., de Girolamo, G., Nakane, H., Mneimneh, Z., Lara, C., de Graaf, R., Scott, K.M., Gureje, O., Stein, D.J., Haro, J.M., Bromet, E.J., Kessler, R.C., Alonso, J., Von Korff, M., 2007. Mental disorders among persons with chronic back or neck pain: results from the World Mental Health Surveys. Pain, June 129, 332-342.
- Elliott, A.M., Smith, B.H., Smith, W.C., Chambers, W.A., 2000. Changes in chronic pain severity over time: the Chronic Pain Grade as a valid measure. Pain 88, 303-308.
- Fielding, R., Wong, W.S., 2008. The prevalence of chronic pain, fatigue, and insomnia in the general population of Hong Kong. Final report to the Health, Welfare and Food Bureau, Government of the Hong Kong Special Administrative Region, China School of Public Health, the University of Hong Kong, Hong Kong.
- Geisser, M.E., Roth, R.S., Robinson, M.E., 1997. Assessing depression among persons with chronic pain using the Center for Epidemiological Studies and the Beck Depression Inventory: A comparative analysis. The Clinical journal of pain 13, 163-170.
- Henderson, S., Andrews, G., Hall, W., 2000. Australia's mental health: An overview of the general population survey. Australian and New Zealand Journal of Psychiatry 34, 197-205.

- Jenkins, R., Bebbington, P., Brugha, t., Farrell, M., Gill, B., Lewis, G., Meltzer, H., Petticrew, M., 1997. The National Psychiatric Morbidity surveys of Great Britian Strategy and methods. Psychological Medicine 27, 765-774.
- Jensen, M.K., Thomsen, A.B., Hojsted, J., 2006. 10-year follow-up of chronic non-malignant pain patients: Opioid use, health related quality of life and health care utilization. European Journal of Pain 10, 423-433.
- Jordanova, V., Wickramesinghe, C., Gerada, C., Prince, M., 2004. Validation of two survey diagnostic interviews among primary care attendees: A comparison of CIS-R and CIDI with SCAN ICD-10 diagnostic categories. Psychological Medicine 34, 1013-1024.
- Lee, Y., Lin, P.-Y., Hsu, S.-T., Cing-Chi, Y., Yang, L.-C., Wen, J.-K., 2008. Comparing the use of the Taiwanese Depression Questionnaire and Beck Depression Inventory for screening depression in patients with chronic pain. Chang Gung Medical Journal 31, 369-377.
- Lewis, G., Pelosi, A.J., Araya, R., Dunn, G., 1992. Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. Psychological medicine 22, 465-486.
- Love, A.W., 1987. Depression in chronic low back pain patients: Diagnostic efficiency of three self-report questionnaires. Journal of Clinical Psychology 43, 84-89.
- Mulrow, C.D., Williams, J.W., Gerety, M.B., Ramirez, G., Montiel, O.M., Kerber, C., 1995. Case-finding instruments for depression in primary care settings. Annal of Internal Medicine 122, 913-921.
- Ohayon, M.M., 2004. Specific characteristics of the pain/depression association in the general population. Journal of Clinical Psychiatry. Special Issue: Recognizing the Physical Symptoms of Depression 65, 5-9.
- Patten, S.B., Williams, J.V., Wang, J., 2006. Mental disorders in a population sample with musculoskeletal disorders. BMC musculoskeletal disorders, Jan 7, 37.
- Preventive Services Task Force, U., 2002a. Screening for depression in adults: A summary of the evidence for the US Preventive Services Task Force. Annals of Internal Medicine 136, 765-776.
- Preventive Services Task Force, U., 2002b. Screening for depression: Recommendations and rationale. Annal of Internal Medicine 136, 760-764.
- Radloff, L.S., 1977. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1, 385-401.
- Reynolds, W.M., Gould, J.W., 1981. A psychometric investigation of the standard and short form Beck Depression Inventory. Journal of consulting and clinical psychology 49, 306-307.
- Shek, D.T., 1991. What does the Chinese version of the Beck Depression Inventory measure in Chinese students--general psychopathology or depression? Journal of Clinical Psychology 47, 381-390.
- Silberstein, S.D., 2001. Shared mechanisms and comorbidities in neurologic and psychiatric disorders. Headache 41 Suppl 1, S11-17.
- Singleton, N., Bumpstead, R., O'Brien, M., Lee, A., Meltzer, H., 2001. Psychiatric morbidity among adults living in private households, 2000. The Stationery Office, London.
- Smith, B.H., Penny, K.I., Purves, A.M., Munro, C., Wilson, B., Grimshaw, J., Chambers, W.A., Smith, W.C., 1997. The Chronic Pain Grade questionnaire: validation and reliability in postal research. Pain 71, 141-147.
- SPSS Inc, 2002. Statistical Package for the Social Sciences. Author, Chicago.

- Tripp, D.A., VanDenKerkhof, E.G., McAlister, M., 2006. Prevalence and determinants of pain and pain-related disability in urban and rural settings in southeastern Ontario. Pain Research & Management 11, 225-233.
- Turk, D.C., Okifuji, A., Scharff, L., 1995. Chronic pain and depression: Role of perceived impact and percieved control in different age cohorts. Pain 61, 93-101.
- Turner, J.A., Romano, J.M., 1984. Self-report screening measures for depression in chronic pain patients. Journal of Clinical Psychology 40, 909-913.
- Von Korff, M., Ormel, J., Keefe, F.J., Dworkin, S.F., 1992. Grading the severity of chronic pain. Pain 50, 133-149.
- Ward, N.G., Bloom, V.L., Friedel, R.O., 1979. The effectiveness of tricyclic antidepressants it the treatment of coexisting pain and depression. Pain 7, 331-341.
- WHO, I.C.i.P.E., 2000. Cross-national comparisons of the prevalence and correlates of mental disorders. Bulletin of the World Health Organization 78, 413-426.
- Whooley, M.A., Avins, A.L., Miranda, J., Browner, W.S., 1997. Case-finding instruments for depression: Two questions are as good as many. Journal of General Internal Medicine 12, 439-445.
- Wong, W.S., Fielding, R., accepted. Chronic pain and psychiatric morbidity. Pain Medicine.
- Ying, Y.W., 1988. Depressive symptomatology among Chinese-Americans as measured by the CES-D. Journal of Clinical Psychology 44, 739-746.
- Zax, M., Takahashi, S., 1967. Cultural influences on response style: Comparisons of Japanese and American college students. Journal of Social Psychology 71, 3-10.
- Zigmond, A.S., Snaith, R.P., 1983. The hospital anxiety and depression scale. Acta Psychiatr Scand 67, 361-370.