

## ORIGINAL ARTICLE

# Measurement properties of the CESD scale among individuals with spinal cord injury

WC Miller<sup>1,2,3</sup>, HA Anton<sup>2,4</sup> and AF Townson<sup>2,4</sup>

<sup>1</sup>Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, British Columbia, Canada; <sup>2</sup>International Collaboration on Repair Discoveries, Vancouver, British Columbia, Canada; <sup>3</sup>GF Strong Rehabilitation Research Lab, Vancouver, British Columbia, Canada and <sup>4</sup>Division of Physical Medicine and Rehabilitation, University of British Columbia, Vancouver, British Columbia, Canada

**Background:** Depression is common after spinal cord injury (SCI), yet it can be difficult and costly to diagnose. Screening tools such as the Center for Epidemiological Studies Depression scale (CESD) can assist with case identification; however, insufficient knowledge of their measurement properties exists to use them in the SCI population.

**Objectives:** To assess the reliability and validity of the CESD-20 and the CESD-10.

**Setting:** Tertiary care centre in Vancouver, British Columbia, Canada.

**Methods:** A 2-week retest study of 47 individuals with traumatic SCI. Subjects  $\geq 19$ , who had their SCI for  $\geq 1$  year and had American Spinal Injury Association Impairment scale ranking of A or B. Short Form-36 (SF-36) subscales and a visual analogue scale for fatigue (VAS-F) were used to assess validity using Pearson's correlations coefficients. Internal consistency was assessed using Cronbach's alpha, retest reliability was assessed using intraclass correlation coefficients (ICCs) and Bland–Altman plots. Normative data are presented based on key demographic and clinical factors.

**Results:** Cronbach's alpha was 0.91 and 0.86 and retest reliability was ICC = 0.87 (95% confidence interval (CI) 0.79–0.93) and ICC = 0.85 (95% CI 0.75–0.92) for the CESD-20 and CESD-10, respectively. Minimal bias was evident based on the Bland–Altman plots. The strongest correlations were with outcomes representing mental health ( $r = -0.71$ ), vitality ( $r = -0.60$ ) and the VAS-F ( $r = 0.57$ ). The weakest correlation was with the physical function score of the SF-36 ( $r = -0.37$ ).

**Conclusion:** The CESD-20 and CESD-10 are quick and easy to use. This study provides evidence in support of the reliability and validity.

*Spinal Cord* (2008) 46, 287–292; doi:10.1038/sj.sc.3102127; published online 2 October 2007

**Keywords:** CESD; validity; reliability; depression; measurement

## Introduction

Adjustment after a spinal cord injury (SCI) presents a formidable challenge. While the primary emphasis on adjustment is physical function, psychological adjustment presents an ancillary challenge. Depression following SCI is common and is the most frequently cited psychological issue in SCI populations. Estimates of prevalence range from 11 to 60%,<sup>1–3</sup> with the prevalence in women being approximately twice that of men.<sup>2</sup> The occurrence of depression is not surprising and has even been considered inevitable<sup>4</sup> and perhaps a natural and necessary part of the acceptance process. However, recent studies have shown that depression after SCI is neither inevitable nor

beneficial, and moreover not everyone who sustains an SCI becomes depressed.<sup>5</sup>

Depression following SCI is concerning because of its negative influence on the rehabilitation process, increased length of hospitalization, decreased longevity, increased rates of suicide, restricted community participation and reduced health and daily functioning are all plausible consequences. Data obtained in earlier studies suggest that many newly injured persons, who meet the criteria for depression, remit within 3 months of onset.<sup>6</sup> Because depression is treatable, early identification and treatment are essential to prevent further problems.

Diagnosing depression is not straight forward and various methodological challenges have constrained its study in the SCI population.<sup>7</sup> One particular issue includes the difficulty defining and operationalizing depression as well as the inconsistent use of diagnostic criteria to identify this condition.<sup>8</sup> The gold standard for identifying depression is

Correspondence: Dr WC Miller, Department of Occupational Science and Occupational Therapy, University of British Columbia, T325-2211 Wesbrook Mall, Vancouver, BC, Canada V6T 2B5.

E-mail: bcmiller@telus.net

Received 11 March 2007; revised 11 August 2007; accepted 11 August 2007; published online 2 October 2007

through the use of structured diagnostic interviews based on the DSM-IV Axis I Disorders criteria. These interviews take considerable time and special expertise. As a result, such detailed interviews are not routinely adopted as a regular clinical practice outside of the context of research.<sup>9</sup> Using a screening tool can facilitate case-finding and moreover provide evidence of outcome after intervention.

Several screening tools exist such as the Beck Depression Inventory and the Zung Self-Rating Depression scale. The 20-item Center for Epidemiological Studies Depression scale (CESD) has been used since the late 1970s.<sup>10</sup> It is recognized for its good psychometric qualities, reliability and validity when used with general population and in primary care patients.<sup>11</sup> It has also proven to be a useful indicator of depression with older adults. Like other self-report measures, it is convenient to use due to its ease and low cost of administration; however, to date, the measurement properties have not been well established for the SCI population. One clinically attractive alternative form of this tool is the 10-item CESD. The purpose of this study was to examine the internal consistency, 2-week retest reliability and the concurrent validity of both the 20- and 10-item CESD in a sample of individuals with traumatic SCI. We also provide normative data for a variety of sample characteristics such as age, sex, motor impairment level and duration of injury.

## Materials and methods

### Design/sample

A 2-week follow-up methodological study design was used to address the study objectives. Acknowledging that there is no ideal retest period, we chose a 2-week period as it would likely eliminate any potential effects of memory while minimizing the potential for change in symptoms of depression.<sup>12</sup>

Community living individuals more than 1-year after traumatic SCI were recruited using advertisements in a disability newsletter, locally distributed posters and from the GF Strong Rehabilitation Research database. The GF Strong Rehabilitation Centre is the primary provider of SCI rehabilitation in British Columbia. Inclusion criteria included age over 19; American Spinal Injury Association (ASIA) impairment level of A or B (indicating no motor function below the sacral levels 4 and 5) SCI; no mental health issues requiring active treatment; and residence within 50 km of the study centre. Individuals were excluded if they were unable to provide informed consent or read English; did not provide complete information; had ventilator-dependent tetraplegia; had active medical conditions such as anemia, hypothyroidism or sleep apnea as identified by physician screen; or experienced a change in health status (new medical conditions or medication changes), activity, caregiver and social situation over the retest interval. We excluded ventilator-dependent individuals with SCI as they make up a small proportion of our accessible population. Therefore, excluding this subgroup would keep the sample homogeneous.

### Protocol

Upon providing consent, subjects were scheduled for two visits to the research laboratory approximately 2 weeks apart. During session one, subject demographics (age, education, marital and work status), SCI factors (date and cause of injury, mobility device use) and all standardized measures were collected. The ASIA assessment was performed by a physician trained in the protocol. The order of the self-report measures completed for the study was randomized. All but four subjects returned to the centre for the second session; those who did not return provided information by phone and e-mail. Subjects were screened for changes in health status (change in medication use, hospitalizations and new diagnoses), activity, caregiver and social situation. The study protocol was approved by the Clinical Research Ethics Review Board of the University of British Columbia.

### Measurement

The 20-item CESD (CESD-20) was originally developed to screen for the frequency of symptoms of depression in the general population. The questions for this tool were generated by selecting items from previously validated tools.<sup>10</sup> Its components cover elements related to depressed mood, feelings of guilt and worthlessness, helplessness and hopelessness, psychomotor retardation, appetite loss and sleep disturbance. Responses capture the frequency of feelings and behaviours over the past 7 days and are rated on a 4-point scale ranging from 0 (rarely or none of the time) to 3 (most or all of the time). Items number 4, 8, 12 and 16 are reverse scored and once corrected an overall score is calculated by summing the scores with higher scores suggesting greater levels of depressive symptoms. Scores of over 15 have been reported to indicate depression in the general population based on DSM III criteria; however, Kuptniratsaikul *et al.*<sup>13</sup> suggested scores over 19 should be used for individuals with SCI. They reported high sensitivity (80%) and reasonable specificity (68%) using this cut score. In a separate study, the internal consistency among a diverse sample of individuals with SCI was  $\alpha=0.88$ .<sup>14</sup> The shortened CESD-10 was created by Andresen *et al.*<sup>15</sup> Less has been reported about the psychometric properties of the shorter version, but Andresen *et al.* reported high agreement with respect to predictive accuracy ( $\kappa=0.97$ ) between cut scores of  $>10$  on the CESD-10 and scores of  $>15$  on the CESD-20. Validity has been reported in the predicted direction with measures of poorer health status ( $r=0.37$ ) and positive affect ( $r=-0.67$ ) in a large sample of well-older adults. Retest correlations in the same sample were adequate ( $r=0.71$ ).

Standardized measures other than the CESD used in this study were included to assess the validity. These tools consisted of a visual analogue scale for fatigue (VAS-F) and the Short Form-36 (SF-36). The VAS-F is a one-item self-report method of capturing global fatigue. The participant marks a point along a 100-mm line indicating the severity of his/her fatigue over the past week. The anchors for the scale are 'absolutely no fatigue' and 'extremely severe fatigue'. It was

included in this study given the moderate association between fatigue and depression.<sup>16</sup>

The SF-36 is the most widely used measure for assessing health status. It was designed for use in clinical practice and research<sup>17</sup> to provide information about the individual's mental and physical health status. For the purposes of this study we used all eight of the subscales that make up the SF-36. These include the mental health, emotional role function, vitality, pain, physical health, physical role function, social role function and general health.

#### Data analysis

To determine the minimum number of subjects needed for the study, we based our sample size calculation on the work of Donner and Eliasziw.<sup>18</sup> Given the relatively short retest period we hypothesized the retest reliability of the CESD using intraclass correlation coefficient (ICC) would be  $>0.80$ . We determined that given sufficient power (0.80) and an alpha of 0.05, 42 subjects would be needed.<sup>18</sup> We over-sampled to compensate for missing data or failure to attend session two.

Descriptive statistics such as means, standard deviations and proportions were used to provide a summary of all variables. In addition, mean CESD values were derived relative to demographic and SCI-related factors. *t*-tests and one-way analysis of variance, for binary and multi-group ( $>2$ ), respectively, were conducted to assess if statistically relevant group differences existed for these factors to provide normative data based on the sample.

Cronbach's alpha was derived to assess reliability related to internal consistency and one-way analysis of variance ICCs (1,1) were calculated to assess the 2-week reliability of the total score and the item-by-item scores of the CESD. The Bland-Altman or limits of agreement method was used to provide a visual assessment of within-test agreement between the baseline and follow-up ratings and to identify any bias that might exist.<sup>19</sup> The difference in CESD scores between the two data collection sessions is plotted against the average for each subject. The mean and standard deviation of the mean difference as well as the true value of the mean using 95% confidence intervals (CIs) were calculated to further assess the existence of bias. Visual inspection of the plot is used to determine if bias is likely. Ideally the data points will be distributed evenly along the 'zero line' if there are no ceiling or floor effects, and an equal number of points should lie above and below if there is no pretest/post-test bias. Finally, the data points should fall within the 95% confidence intervals.

Analysis of validity was performed using Pearson's product moment correlation. We hypothesized the CESD would have a statistically significant positive correlation with the VAS-F and a negative correlation with each of the eight SF-36 subscales. We anticipated moderate or higher correlations ( $r > 0.5$ ) with scores that represented mental health attributes such as the vitality, mental health and emotional role limitation subscales of the SF-36 and the VAS-F. Weaker correlations were expected between the CESD scores and the physical and general health subscales of the SF-36. The level

of statistical significance was set at  $P < 0.05$ . All analyses were completed using SPSS version 14.1.

## Results

Fifty-five individuals were recruited; however, five did not provide information on the CESD at session two, two reported a major change during the retest period and one subject exceeded the 2-week follow-up collection period by 2 weeks. The mean age of the remaining 47, primarily male (64%) subjects, was 40.6 ( $\pm 12.6$ ) years. The range in years from the time of lesion was 1–41 years, the majority had an ASI A classification (68%) and had their injury related to traumatic reasons (87%). Other sample demographic and SCI-related information is presented in Table 1.

The mean CESD-20 score for the sample was 15.2 (range 0–42) with 39% of the sample scoring over 15 and 30% scoring over 19 and the mean CESD-10 score was 9.7 (range 0–25) with 38.3% scoring over 10. The only significant CESD mean difference based on demographic or clinical factors was for those who were working or attending school versus those who were not (Table 1). Although not statistically significant fewer symptoms of depression were evident the longer one had a SCI.

Internal consistency measured using Cronbach's alpha was 0.91 for the CESD-20 and 0.86 for the CESD-10. Item-wise deletion revealed that the alpha did not change by more than 0.05 with the exclusion of any item. The total score test-retest reliability was ICC = 0.87 (95% CI 0.79–0.93) for the CESD-20 and ICC = 0.85 for the CESD-10 (95% CI 0.75–0.92). The item-by-item reliability (Table 2) revealed that the lowest reliability was for item 4 (ICC = 0.30) and highest for item 3 (ICC = 0.73).

The Bland-Altman limits of agreement plots (Figures 1a and b) indicate that an equal number of subjects scored above and below zero, suggesting that there was minimal bias towards scoring higher or lower during the baseline or follow-up stages. The distribution of the points tended to cluster closer to the left-hand side of the plot. However  $<15\%$  of the sample scored at either extreme of the scale suggesting there is likely no floor or ceiling effect.<sup>19,20</sup> Three of 47 data points were found to lie outside of the two standard deviations of the mean difference, suggesting that the error in reliability readings for the test-retest reliability may be statistically significant.<sup>19</sup> A *t*-test of the mean values demonstrated there was no statistical difference for the CESD-20 (mean difference =  $-1.3$ ; 95% CI =  $-2.76, 0.24$ ;  $t = -1.70$ ;  $P = 0.10$ ) or the CESD-10 (mean difference =  $-0.22$ ; 95% CI =  $-1.21, 0.78$ ;  $t = -0.44$ ;  $P = 0.66$ ) at time 1 and time 2. The inclusion of 0 in the 95% CI suggests minimal bias between the measurement sessions.

The expected associations between the CESD and the other measures were observed in terms of the hypothesized direction and magnitudes of the correlations (Table 3). The strongest correlation was with the mental health subscale with approximately 50% shared variance followed by the vitality, fatigue, general health and role emotional subscale scores, all with approximately 21–36% shared variance.

**Table 1** Sample descriptive statistics and mean CESD-20 and CESD-10 values by sample factors ( $N=47$ )

	Sample descriptives	Mean (s.d.) CESD-20	Mean (s.d.) CESD-10
<i>Demographic factors</i>			
Mean age (s.d.)	40.6 (12.6)		
Age tertiles			
19–36		16.8 (12.8)	10.6 (7.2)
36.1–46		14.4 (8.4)	8.9 (4.9)
> 46		14.2 (8.8)	9.5 (5.9)
Sex (%)			
Male	64.0	14.8 (9.9)	9.6 (6.4)
Female	36.0	15.9 (10.1)	9.9 (5.5)
Education (%)			
High school or less	34.0	14.3 (10.1)	9.4 (6.5)
Post-secondary	66.0	15.7 (9.9)	9.9 (5.9)
Marital status (%)			
Single	72	15.8 (10.2)	9.9 (6.3)
Married/common law	28.0	13.6 (9.0)	9.4 (5.6)
Employment status (%)			
Unemployed	75.0	11.0 (6.1)	10.8 (6.4)
Employed/student	26.0	16.7 (10.5)*	6.7 (3.4)*
<i>SCI-related factors</i>			
ASIA level (%)			
ASI A	62	15.8 (9.4)	9.9 (5.8)
ASI B	38	14.3 (10.9)	9.5 (6.6)
Mean years since injury	15.2 (11.7)		
Tertiles			
1–6.0		17.2 (14.9)	10.9 (8.6)
6.1–20		15.3 (5.7)	9.9 (4.1)
>20		13.5 (8.1)	8.6 (4.8)

Abbreviations: ASIA, American Spinal Injury Association; CESD, Center for Epidemiological Studies Depression; s.d., standard deviation.

\* $P < 0.05$  for difference between groups.

**Table 2** Item wise and total ICC values for CESD

Item number & brief description	ICC	CI 95%
(1) I was bothered by things	0.32	0.04–0.55
(2) My appetite was poor	0.51	0.27–0.69
(3) I could not shake off the blues	0.73	0.56–0.84
(4) I felt that I was just as good as other people	0.11	0.01–0.38
(5) <b>Trouble keeping my mind on what I was doing</b>	0.62	0.41–0.77
(6) <b>I felt depressed</b>	0.67	0.48–0.80
(7) <b>Everything I did was an effort</b>	0.59	0.37–0.75
(8) <b>I felt hopeful about the future</b>	0.59	0.37–0.75
(9) I thought my life had been a failure	0.28	–0.03–0.52
(10) <b>I felt fearful</b>	0.47	0.22–0.67
(11) <b>My sleep was restless</b>	0.64	0.43–0.78
(12) <b>I was happy</b>	0.52	0.28–0.70
(13) I talked less than usual	0.64	0.43–0.78
(14) <b>I felt lonely</b>	0.55	0.32–0.72
(15) People were unfriendly	0.43	0.17–0.64
(16) I enjoyed life	0.63	0.41–0.77
(17) I had crying spells	0.54	0.30–0.71
(18) I felt sad	0.62	0.41–0.77
(19) I felt that people disliked me	0.71	0.53–0.83
(20) <b>I could not get “going”</b>	0.68	0.49–0.81

Abbreviations: CESD, Center for Epidemiological Studies Depression scale; ICC, intraclass correlation coefficients.

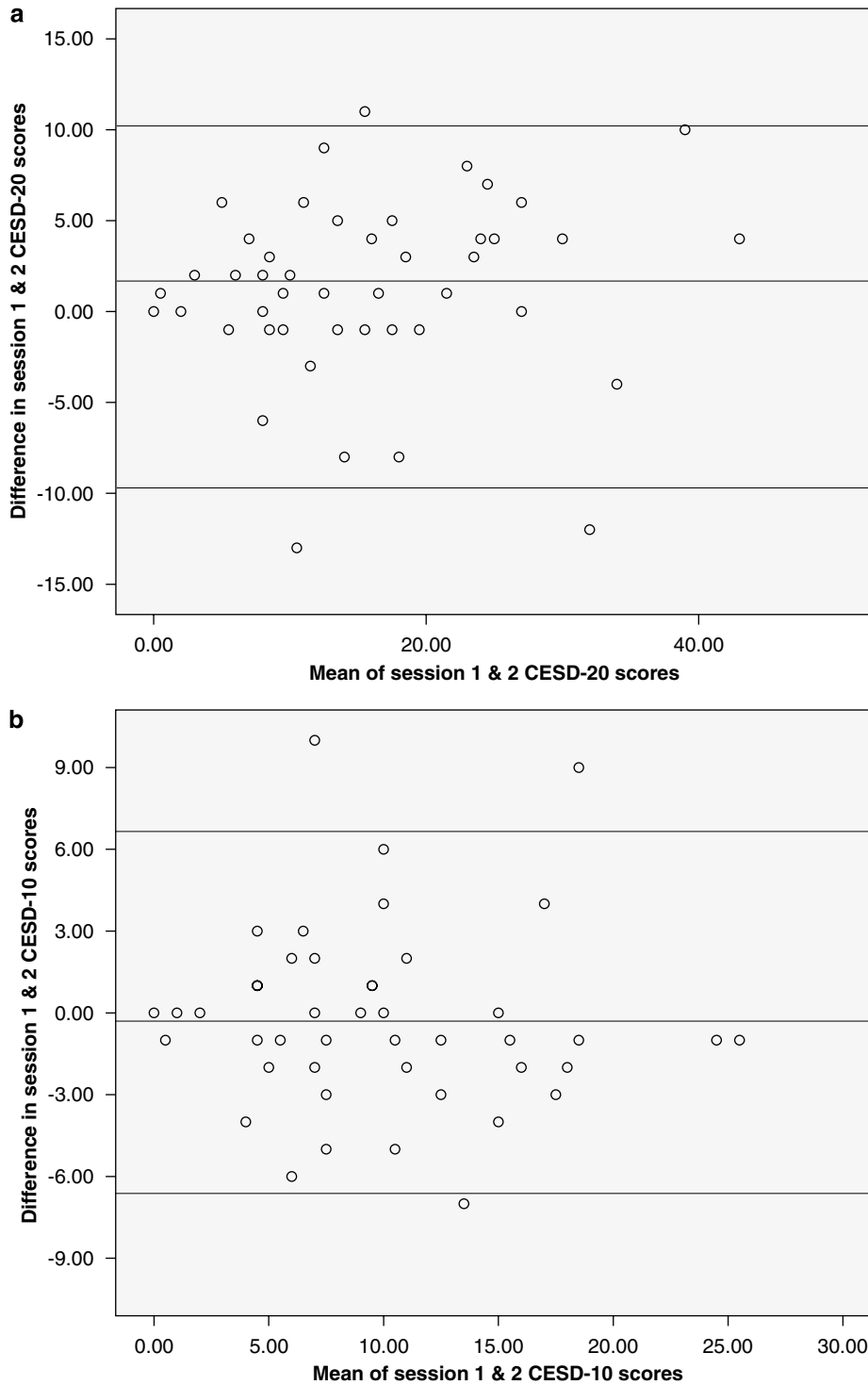
Texts in bold indicate those items included in the CESD-10.

The weakest correlation was with the physical functioning and pain subscales.

## Discussion

The results of this study provide support for the reliability and validity of the original CESD-20 and the CESD-10. The internal consistency of the CESD-20 was marginally better than the CESD-10, which can be expected simply based on the difference in the number of items included; however, the retest reliability was virtually the same. Both the internal consistency and the retest reliability coefficients exceed the criteria Andresen *et al.*<sup>15</sup> advocate as indicative of a top grade.

The Pearson's correlations between the validity measures and both versions of the CESD were in the same direction and of similar magnitude, although the coefficients were slightly larger for the CESD-10 except for the association with the SF-36 mental health score. The magnitude of the validity coefficients was similar to those reported in different populations. For instance, the mental health subscale score was related to ( $r = -0.76$ ) the CESD in a sample of individuals



**Figure 1** (a) Bland–Altman limits of agreement plot of CESD-20. (b) Bland–Altman limits of agreement plot of CESD-10.

with rheumatoid arthritis.<sup>21</sup> Previous studies in cancer survivors,<sup>22</sup> individuals with sarcoidosis<sup>23</sup> and family caregivers of stroke survivors<sup>24</sup> also reported stronger correlation of the mental health component scores to CESD than the physical health scores. The consistency of these findings with those of our study supports the notion that the CESD is valid for use with the SCI population.

Limitations of our study include the generalizability of our study results. Our sample did not include individuals in the first year after SCI or those with motor incomplete SCI. Moreover, the sample was well educated. A larger sample would have enabled us to assess the factor structure of the tool for the target population. Our retest period was 2 weeks and although we are confident that is long enough to reduce

**Table 3** Pearson's correlations of the CESD-20 and CESD-10 with other measures

Measure	CESD-20	CESD-10
	r	r
VAS-F	0.52	0.57
SF-36 mental health	0.75*	0.71*
SF-36 emotional role function	0.55*	0.56*
SF-36 vitality	0.54*	0.60*
SF-36 pain	0.27*	0.38*
SF-36 social role function	0.37*	0.42*
SF-36 physical function	0.34*	0.37*
SF-36 physical role function	0.40*	0.49*
SF-36 general health	0.57*	0.60*

VAS-F, visual analogue scale of fatigue.

\* $P < 0.05$  was considered to be significant.

any influence of learning on the subjects' responses,<sup>12</sup> it is plausible that a longer period may have eliminated this potential source of bias. A number of variables were used to assess the construct validity of the CESD by determining how well it converged with a measure of fatigue and attributes of the SF-36. We recognize that using a single item measure of fatigue is less than ideal; however, we are unaware of any fatigue measures that have been validated with the SCI population. Therefore, when considered in combination with an effort to minimize burden on the sample, we elected to use the VAS-F. Finally, screening tools like the CESD should not be used to diagnose depression; however, they do provide valuable information that may alert clinicians to conduct further evaluation.

## Conclusion

Depression after SCI can restrict participation in life and may predispose individuals to medical complications. Moreover on the basis of our clinical experience, depression may interfere with participation in inpatient and outpatient rehabilitation. The challenge for clinicians is to distinguish those normal emotional responses to SCI from clinically significant depression that requires treatment. The CESD-20 and CESD-10 are quick and easy to use and have the advantage of published normative data from studies of the general population. This study provides evidence in support of the reliability and validity of these tools as screening tools for depression and normative data for individuals with SCI.

## Acknowledgements

We thank the subjects who participated in this study as well as other members of the Fatigue and Function in Neurologic and Neuromuscular Disorders (F2N2) study group who provided encouragement and support throughout the project. This work was financially supported by the Rick Hansen Man in Motion Research Foundation. Additional support has been provided by the Michael Smith Foundation for Health Research Team Planning Award with matching funds from the Vancouver Coastal Health Research Institute. Dr Miller receives salary support from the Canadian Institutes of Health Research.

## References

- Bombardier CH, Richards JS, Krause JS, Tulskey D, Tate DG. Symptoms of major depression in people with spinal cord injury: implications for screening. *Arch Phys Med Rehabil* 2004; **85**: 1749–1756.
- Consortium for Spinal Cord Medicine. [http://www.pva.org/site/PageServer?pagename=pubs\\_generalpubs](http://www.pva.org/site/PageServer?pagename=pubs_generalpubs). Paralyzed Veterans of America: Washington, DC, 1998.
- Kennedy P, Rogers B. Reported quality of life of people with spinal cord injuries: a longitudinal analysis of the first 6 months post-discharge. *Spinal Cord* 2000; **38**: 498–503.
- Hohmann GW. Psychological aspects of treatment and rehabilitation of the spinal cord injured person. *Clin Orthop Relat Res* 1975; **112**: 81–88.
- Kemp BJ, Kahan JS, Krause JS, Adkins RH, Nava G. Treatment of major depression in individuals with spinal cord injury. *J Spinal Cord Med* 2004; **27**: 22–28.
- Kishi Y, Robinson RG, Forrester AW. Prospective longitudinal study of depression following spinal cord injury. *J Neuropsychiatry Clin Neurosci* 1994; **6**: 237–244.
- Elliott TR, Frank RG. Depression following spinal cord injury. *Arch Phys Med Rehabil* 1996; **77**: 816–823.
- Snaith RP. Defining 'depression'. *Am J Psychiatry* 1987; **144**: 828–829.
- Farvolden P, McBride C, Bagby RM, Ravitz P. A web-based screening instrument for depression and anxiety disorders in primary care. *J Med Internet Res* 2003; **5**: e23.
- Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977; **1**: 385–401.
- Fave GA, Pilowsky I, Perfederici A, Bernardi M, Pathak D. Depressive symptoms and abnormal illness behaviour in primary care patients. *Gen Hosp Psychiatry* 1982; **4**: 171–178.
- Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*, 2nd ed. Oxford University: Oxford, New York, 1995.
- Kuptniratsaikul V, Chulakadabba S, Ratanavijitrasil S. An instrument for assessment of depression among spinal cord injury patients: comparison between the CES-D and TDI. *J Med Assoc Thai* 2002; **85**: 978–983.
- Lachapelle DL, Alfano DP. Revised Neurobehavioral Scales of the MMPI: sensitivity and specificity in traumatic brain injury. *Appl Neuropsychol* 2005; **12**: 143–150.
- Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *Am J Prev Med* 1994; **10**: 77–84.
- Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The Fatigue Severity Scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989; **46**: 1121–1123.
- Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**: 473–483.
- Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med* 1987; **6**: 441–448.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**: 307–310.
- McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995; **4**: 293–307.
- Escalante A, del Rincon I, Mulrow CD. Symptoms of depression and psychological distress among Hispanics with rheumatoid arthritis. *Arthritis Care Res* 2000; **13**: 156–167.
- Sarna L, Padilla G, Holmes C, Tashkin D, Brecht ML, Evangelista L. Quality of life of long-term survivors of non-small-cell lung cancer. *J Clin Oncol* 2002; **20**: 2920–2929.
- Cox CE, Donohue JF, Brown CD, Kataria YP, Judson MA. Health-related quality of life of persons with sarcoidosis. *Chest* 2004; **125**: 997–1004.
- Grant JS, Weaver M, Elliott TR, Bartolucci AA, Giger JN. Family caregivers of stroke survivors: characteristics of caregivers at risk for depression. *Rehabil Psychol* 2004; **49**: 172–179.