International Journal of Nursing Studies 97 (2019) 14–20

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns

Psychometric properties of the Chinese version of the Center for Epidemiologic Studies Depression Scale-Revised in patients with cancer: A cross-sectional study

Jie Li[a], Fengling Zhao[b], Huayu Bai[a], Pingzhen Lin[a], Di Shao[a], Jiwei Sun[a], Huihui Zhang[a], Fenglin Cao[a,*]

[a] School of Nursing, Shandong University, Jinan, Shandong Province, 250012, China
[b] Department of Mental Health Care, Shandong Cancer Hospital, Jinan, Shandong Province, 250117, China

ARTICLE INFO

Article history:
Received 19 December 2018
Received in revised form 24 March 2019
Accepted 13 April 2019

Keywords:
Cancer
Chinese version
COSMIN checklist
Depression
Depressive symptoms
Oncology
Psychometrics
Reliability
Validity

ABSTRACT

Background: Depressive symptoms are common in patients with cancer and more prevalent in Chinese patients. The Center for Epidemiologic Studies Depression Scale is one of the most widely used self-report scales to assess depressive symptoms in both community and hospitalized samples. A revised Center for Epidemiologic Studies Depression Scale has been created, but the evidence on psychometric properties is limited.
Objectives: To develop the Chinese version of the scale, and to examine the cross-cultural validity, structural validity, construct validity, internal consistency, test-retest reliability, measurement error, responsiveness, and floor/ceiling effect of the scale among patients with cancer according to the recommendation in the consensus-based standards for the selection of health status measurement instruments checklist.
Design: A cross-sectional survey design with 33 participants (approximately 10%) completing the follow-up survey for evaluating the test-retest reliability.
Settings: Randomly selected eight wards of an oncology hospital in China.
Participants: Of the 595 patients we approached in the randomly selected eight wards, 310 gave their informed consent and completed the survey.
Methods: The Chinese version of the Center for Epidemiologic Studies Depression Scale-Revised was developed by four researchers (two translators and two reviewers) who were proficient in both English and Chinese. Participants completed the scale and the depression module of the Patient Health Questionnaire. Principal components analyses, Spearman's correlation, the Mann-Whitney *U* test, Cronbach's alpha, and the intraclass correlation coefficient were used.
Results: The cross-culture validity was excellent based on the consensus-based standards for the selection of the health status measurement instruments checklist. A two-factor structure was determined: somatic symptoms and affective-cognitive symptoms. The sufficient construct validity was supported by that the score of the Center for Epidemiologic Studies Depression Scale-Revised was strongly correlated with the depression module of the Patient Health Questionnaire score ($rho = 0.73$, $p < 0.001$) and patients who had received chemotherapy ($p = 0.002$)/radiotherapy ($p = 0.035$) reported higher scores of depression than those who have not. The Cronbach's alphas of the total scale and subscales ranged from 0.82–0.88. The test-retest reliability was sufficient (intraclass correlation coefficient = 0.73–0.81, all $p < 0.001$) for total scale and subscales. The smallest detectable change was 2.98 and the responsiveness was adequate, with no floor/ceiling effect.
Conclusions: This study supports the Chinese version of the Center for Epidemiologic Studies Depression Scale-Revised as a valid and reliable measurement of depressive symptoms in patients with cancer.

What is already known about the topic?

- Depression is common in patients with cancer and has negative impact on patients.
- The Center for Epidemiologic Studies Depression Scale (CESD) is one of the most widely used self-report scales to assess depressive symptoms in patients with cancer.

* Corresponding author at: School of Nursing, Shandong University, Jinan, Shandong Province, 250012, China.
E-mail address: caofenglin2008@126.com (F. Cao).

https://doi.org/10.1016/j.ijnurstu.2019.04.008
0020-7489/

- The revised CESD (CESD-R) have been developed, but the evidence of its psychometric properties is limited.

What this paper adds

- The Chinese version of the CESD-R is valid, as shown by the excellent cross-cultural validity, structural validity, and construct validity.
- It is demonstrated to be reliable by the excellent internal consistency and good test-retest reliability.
- The responsiveness is adequate and there are no floor or ceiling effects.

1. Introduction

1.1. Background

The incidence of cancer is increasing around the world. It is estimated that by 2035 about 24 million new cases of cancer will be diagnosed annually worldwide. In China, the estimated annual incidence is expected to reach 5.5 million, not including the Chinese-speaking population in other countries (Ferlay et al., 2013), which indicates that the new cases of cancer who are Chinese-speaking around the world are more than 5.5 million per year.

Cancer not only impairs physical health but also leads to psychological problems such as depression. Among patients with cancer in developed countries, the prevalence of depression assessed by self-report measures ranges from 12.9%–27.0% (Krebber et al., 2014; Linden et al., 2012; Mols et al., 2018). Studies reporting the prevalence of depression among Chinese patients with cancer have found the prevalence to be 32.4%–66.7% (Bao et al., 2017; Hong and Tian, 2014; Liu et al., 2017), which is higher than the prevalence of depression in patients in other developed countries. Depression can lead to increased healthcare visits and prolonged hospitalization (Mausbach and Irwin, 2017), an adverse impact on compliance with cancer therapy (Berry et al., 2015), a diminished quality of life (Mols et al., 2018), and even reduced survival time (Sullivan et al., 2016) in patients with cancer. Thus, considering the large number of cancer patients struggling with depression and its negative impact, it is necessary to develop a valid instrument to assess the depressive symptoms of patients with cancer. Moreover, in order to improve the comparability of depression in cancer patients across populations, a widely used instrument is necessary.

The Center for Epidemiologic Studies Depression Scale (CESD) is one of the most widely used self-report scales with good psychometric properties to assess depressive symptoms in the general population (Radloff, 1977; Roberts and Vernon, 1983) and in hospital-based samples (Janssen et al., 2019; Lynn et al., 2019; Moon et al., 2017), including patients with cancer (Chhabria and Carnaby, 2017; Muzzatti et al., 2018; Schroevers et al., 2000; Sullivan et al., 2016; Wakefield et al., 2015). In 2004, William W. Eaton et al revised the CESD and created the revised version (CESD-R) (Eaton et al., 2004). The CESD-R changed some items (e.g., eliminating items measuring positive affect and adding items measuring anhedonia, psychomotor function, and suicidal ideation) and one response category (adding "nearly every day for two weeks") to improve the reliability and address the updating of the diagnostic criteria of depression in the Fourth Edition of the Diagnostic and Statistical Manual of Mental Disorders (Eaton et al., 2004). The CESD-R has been tested to show good reliability and validity among the general population in the United States (Eaton et al., 2004; Van Dam and Earleywine, 2011). Validation studies of the CESD-R in the general populations have been conducted for the English, Polish (Koziara, 2016) and Korean (Lee et al., 2016) versions. However, the evidence is limited and the psychometric properties of the Chinese version have not yet been reported.

1.2. Objectives

Depressive symptoms in patients with cancer are prevalent and even more common in Chinese patients with cancer (Bao et al., 2017; Hong and Tian, 2014; Krebber et al., 2014; Linden et al., 2012; Liu et al., 2017; Mols et al., 2018). A psychometrically valid and widely used instrument for evaluating the symptoms is needed. The CESD-R made up for some shortcomings of the old version and addressed the new diagnostic criteria of depression. However, there is currently limited evidence on psychometric properties of the CESD-R and the results were not reported according to a standard guideline, such as the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist (Mokkink et al., 2010a,2010b). The COSMIN was developed in 2010 by consulting international experts on psychology, epidemiology, statistics and clinical medicine in a Delphi study, and can be used as a guideline of study on measurement properties of health-related patient-reported outcomes (Mokkink et al., 2010a, 2010b). Thus, this study aimed to develop the Chinese version of the CESD-R and test its validity (cross-cultural validity, structural validity, and construct validity), reliability (internal consistency and test-retest reliability), responsiveness validity, and floor/ceiling effect according to the COSMIN checklist in patients with cancer.

2. Methods

2.1. Study design and setting

Eight randomly selected wards from an oncology hospital in East China from April to July 2016 were included in this cross-sectional study. These were selected from a pool of 39 wards in the hospital: ten wards of radiotherapy, ten wards of medical oncology, 15 wards of surgical oncology, and four wards of gynecological oncology. Two wards of radiotherapy, two wards of medical oncology, three wards of surgical oncology, and one ward of gynecological oncology were randomly selected in proportion (1:5).

2.2. Sample

Hospitalized patients of these eight wards were recruited. The inclusion criteria included patients who were (1) at least 18 years old, (2) diagnosed with cancer, (3) undergoing cancer treatments (surgery, chemotherapy, or radiotherapy), and (4) had adequate reading and writing ability to finish the questionnaire. We excluded participants with other physical or psychological disturbances that would obstruct their participation. This study was approved by the Ethics Committee of Shandong University School of Nursing (R-026) and all participants gave their informed consent. This study was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.3. Developing the Chinese version of the CESD-R

The CESD-R (Eaton et al., 2004), which is an updated version of the CESD (Radloff, 1977), includes 20 items and measures nine symptoms of depression (dysphoria, anhedonia, appetite, sleep, concentration, worthlessness, fatigue, agitation, and suicidal

ideation) as defined by the diagnostic criteria for depression in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (APA, 1994) and fifth edition (APA, 2013). In the CESD-R, each item is rated on a 5-point Likert scale: 0 = "not at all or less than one day in the last week," 1 = "1–2 days in the last week," 2 = "3–4 days in the last week," 3 = "5–7 days in the last week," and 4 = "nearly every day for the last two weeks." Adding the score for each item produces the total score and it ranges from 0 to 80, with higher scores indicating more severe depressive symptoms. The CESD-R was supported to have high internal consistency (0.92–0.93) and good construct validity in the general population of the United States (Van Dam and Earleywine, 2011).

We obtained the English version of the CESD-R from the official website (http://cesd-r.com) (Eaton et al., 2012) and acquired permission from Professor Eaton to use and translate it into Chinese. The scale was translated into Chinese by two researchers of health psychology (focusing on psycho-oncology) who are proficient in both English and Chinese, following the guidelines prescribed in the Translation and Cultural Adaptation-Principles of Good Practice (Wild et al., 2005). These two researchers independently translated it into Chinese and compared the two translations. After discussion, they could not agree on the translation of Item 4 ("I felt depressed.") and Item 7 ("I could not get going."). A professor (PhD, focusing on psychology, and having developed four scales of the Chinese version from the original English version) who is also proficient in both English and Chinese reviewed the translation and discussed Items 4 and 7 with the two translators to obtain agreement. Then, the two translators respectively translated the Chinese scale back into English and compared their translations with the original one to test whether the Chinese scale reflected the same meaning as the original CESD-R. The professor reviewed their translations and reconciled the differences to develop a single Chinese version of the CESD-R. Then, five oncology nurses and seven postgraduate nursing students with research area specialty of health psychology completed the scale, which was followed by a discussion to evaluate the understandability, interpretation, and cultural relevance of the scale. Specifically, the alternative wording of Items 4 and 7 was discussed and a single conclusion was reached. Finally, the translation was reviewed by an assistant professor (PhD, focusing on child and adolescent psychology) who has study experience in both China and the USA and finalized for use in this study.

2.4. Data collection

Eight graduate students were trained to conduct the investigation. Patients who were hospitalized for cancer treatment were approached and assessed for eligibility. Eligible patients were informed about the study, and patients who gave written informed consent were asked to complete a self-reported questionnaire. In order to determine the test-retest reliability, 33 patients returning to the hospital for further treatment agreed to complete the CESD-R about three weeks after the baseline assessment.

2.5. Methodological testing and statistical analysis

All data analyses were performed using SPSS Version 24.0 (IBM Corp., Armonk, NY, USA). Two-tailed significance tests were performed with a p value of 0.05 as the significance level. The distribution of the CESD-R score was not normal based on the Shapiro-Wilk test, so the Spearman's correlation analysis and the Mann-Whitney U test were used. Descriptive statistical analyses were used to report the mean ± standard deviation or medians (25^{th} percentile, 75^{th} percentile) of the continuous variables and the frequency (percentage) of the categorical variables. We evaluated cross-cultural validity, structural validity, construct validity, internal consistency, test-retest reliability, responsiveness, and floor/ceiling effect according to the recommendation in the COSMIN checklist (Mokkink et al., 2010a, 2010b).

2.5.1. Cross-cultural validity

In this study, the Chinese version of the CESD-R was developed from the original English version. Therefore it is important to evaluate the cross-cultural validity, which refers to the extent to which the description of the items on the translated scale can adequately reflect the items of the original scale (Mokkink et al., 2010a, 2010b). The cross-cultural validity was measured by the Cross-cultural validity of the "COSMIN checklist with 4-point scale" (Terwee et al., 2012).

2.5.2. Structural validity

Structural validity is defined as the extent to which the scores of a measurement are an adequate reflection of the dimensionality of the expected construct (Mokkink et al., 2010a, 2010b). The principal components analysis (PCA) was conducted to test the structural validity. According to the COSMIN checklist, it should be conducted before testing the internal consistency for the purpose of checking whether the scale has unidimensionality (Mokkink et al., 2010a, 2010b). The promax rotation and Kaiser's criteria (eigenvalue > 1) were used for factor extraction. To determine the factorial structure more accurately, a parallel analysis (Horn, 1965) and Velicer's minimum average partial (MAP) test (Velicer, 1976; Velicer et al., 2000) were conducted (O'Connor, 2000).

2.5.3. Construct validity (Hypotheses-testing)

Construct validity refers to the extent to which a measure evaluates the intended concept (Mokkink et al., 2010a, 2010b). The depression module of the Patient Health Questionnaire (PHQ-9), a brief self-reported diagnostic and severity measure of depression (Kroenke et al., 2001), was used for assessing construct validity. It was developed according to the diagnostic criteria of depression in the Fourth Edition of the Diagnostic and Statistical Manual of Mental Disorders (APA, 1994), as was the construct of the CESD-R. The PHQ-9 has been supported as a valid and reliable instrument for use in the Chinese general population and medical populations, including with cancer patients (Chen et al., 2010; Sun et al., 2017; Zhang et al., 2016). We tested the construct validity using the Spearman's coefficient between the scores of CESD-R and PHQ-9. We hypothesized that the scores of CESD-R and PHQ-9 were at least moderately correlated (moderate: 0.40−0.59, strong: 0.60−0.79). We also hypothesized that, consistent with prior research (Bower et al., 2011; Gray et al., 2014; Torres et al., 2013), patients who received chemotherapy and radiotherapy would report higher levels of depression. Mann-Whitney U test was used to compare the differences in the CESD-R scores among patients who received or did not receive chemotherapy/radiotherapy.

2.5.4. Internal consistency

Internal consistency is defined as the degree to which each item correlated with each other (Mokkink et al., 2010a, 2010b). A Cronbach's alpha coefficient (Cronbach, 1951), the most widely used index of internal consistency, was calculated to assess the internal consistency of the CESD-R. A Cronbach's alpha of 0.7 and above represents acceptable consistency (Terwee et al., 2007).

2.5.5. Test-retest reliability

Test-retest reliability is defined as the degree to which the results of the measurement are consistent over time (Mokkink et al., 2010a, 2010b). It was assessed using the two-way mixed effects model of the intraclass correlation coefficient (ICC) with an absolute agreement definition (Weir, 2005). The preferable value

of ICC is 0.7 or larger (Mokkink et al., 2010a, 2010b). The test-retest survey was conducted by the same rater as the baseline survey three weeks after the baseline. We assumed that the level of depression would be stable during this time period, when patients were free of active treatment.

2.5.6. Measurement error

Measurement error refers to the systematic and random error of a score of the measurement which is not the reflection of the true differences of measured construct (Mokkink et al., 2010a). The standard error of measurement (SEM) and the smallest detectable change (SDC) was calculated. The SEM indicates the standard deviation of measures within subject and was calculated by the square root of the error variance (Terwee et al., 2007). The SDC was calculated by ($1.96 \times \sqrt{2} \times \text{SEM} \div \sqrt{n}$), which indicates the smallest detectable changes in group level (Terwee et al., 2012).

2.5.7. Responsiveness and floor/ceiling effect

Responsiveness refers to the ability of a measurement to detect change of the construct over time (Mokkink et al., 2010a, 2010b). It was tested by the comparison of SDC and the minimal important change (MIC). A smaller SDC than MIC reflected a sufficient responsiveness (Terwee et al., 2012). The MIC of CESD was 10 reported by Busch et al (Busch et al., 2011). We presumed that the MIC of CESD-R was at least 10 since the possible range of CESD-R (0–80) was large than that of the CESD (0–60). Floor effects were evaluated by examining the percentage of the respondents that achieved the lowest possible scores, and ceiling effects were evaluated by examining the percentage of respondents that reached the highest possible score. A percentage over 15% indicates the existence of a floor or ceiling effect (Terwee et al., 2007).

3. Results

3.1. Sample characteristics

Of the 595 patients we approached, 423 were interested in the study and 327 were eligible. A total of 310 gave their informed consent and completed the questionnaires. The mean age of the patients was 49.68 years (standard deviation = 10.87) and ages ranged from 23 to 82. The majority of the patients were female (74.5%), married (95.5%), and had less than a high school education (74.5%). There were 157 participants (50.6%) who had a family income above RMB 2000 and 150 participants (48.4%) who were employed. Of the participants included in this study, 169 had breast cancer (54.5%), 49 had gastrointestinal cancer (15.8%), 45 had lung cancer (14.5%), 30 had gynecologic cancer (9.7%), and 17 had head and neck cancer (5.5%). There were 146 patients (47.1%) who received surgery, 166 (53.5%) who received chemotherapy, and 50 (16.1%) who received radiotherapy.

3.2. Cross-cultural validity

The 310 participants answered all the 20 items of the CESD-R. The sample size met the requirement of excellent in the COSMIN check list (140) (Terwee et al., 2012). The process of translation met the requirements of excellent in the checklist as well. The cross-cultural validity rated excellent according to the "COSMIN checklist with 4-point scale" except for number 11 about pre-test ("Was the sample used in the pre-test adequately described?"). At the beginning of the survey, we asked the patients if there were some items they could not understand. None of the first 30 patients reported confusion or incomprehension about the items of the CESD-R so we continued to conduct the survey without an independent pre-test.

3.3. Structural validity

The Bartlett sphericity test indicated that the sample was adequate for factor analysis (χ^2 = 2471.889, df = 190, $p < 0.001$; Kaiser-Meyer-Olkin = 0.846). The PCA produced a model of five factors with eigenvalues over 1.00 and explained 61.6% of the variance. The parallel analysis suggested that three roots with eigenvalues larger than the mean eigenvalues were obtained by chance. Velicer's MAP test suggested that there were three underlying components according to the original MAP test (Velicer, 1976) and two factors according to the revised MAP test (Velicer et al., 2000). The PCA with promax rotation was conducted four times, setting the number of factors extracted at two, three, four, and five. In the four- and five-factor solutions, there were two factors with only two items, which suggested a weak and unstable factor (Costello and Osborne, 2005). In the five-factor results, there was one factor related to appetite and sleep disturbances (Items 1, 5, 18, and 19), but it excluded another sleep-related item (Item 11). In the four-factor solution, there was one factor related to fatigue, appetite, and sleep (Items 7, 11, 12, 16, 18, and 19), but it excluded Item 1 ("My appetite was poor."). Item 5 ("My sleep was restless.") was cross-loaded in this factor (factor loading = 0.38) and another factor related to negative mood and concentration impairment (factor loading = 0.39). In the three-factor solution, Item 1, which measured appetite, loaded highest in a factor related to negative mood and concentration impairment, but this factor excluded Item 10, which measures anhedonia. The two-factor solution was both theoretically and psychometrically consistent, suggesting that there is one factor related to somatic symptoms (including appetite, sleep, fatigue, and movement) and another factor related to affective-cognitive symptoms (including sadness, anhedonia, concentration, worthlessness, and suicidal ideation) (see Table 1).

3.4. Construct validity (Hypotheses-testing)

There was a strong positive correlation between the total CESD-R and PHQ-9 scores (rho = 0.73, $p < 0.001$). The patients who had received chemotherapy/radiotherapy reported significantly higher scores of the total symptoms, somatic symptoms, and affective-cognitive symptoms than those had not (see Table 2).

3.5. Internal consistency

The internal consistency analysis showed a Cronbach's alpha of 0.88 for the CESD-R (0.82 for the somatic factor and 0.84 for the affective-cognitive factor). These indicate high internal consistency. In order to confirm that no single item could decrease the internal consistency, we calculated the Cronbach's alpha coefficients with each item removed. The coefficients were 0.87–0.88 for the total CESD-R, 0.78–0.81 for the somatic subscale, and 0.81–0.84 for the affective-cognitive subscale, indicating that no single item significantly decreased the consistency.

3.6. Test-retest reliability

A total of 33 patients completed the follow-up assessment for the test-retest reliability analysis. The ICC value for the CESD-R total score was 0.81 (95% CI: 0.62–0.91, $p < 0.001$); somatic factor was 0.80 (95% CI: 0.60–0.90, $p < 0.001$); and affective-cognitive symptoms was 0.73 (95% CI: 0.46–0.87, $p < 0.001$).

3.7. Measurement error

The SEM was 6.17 and the SDC was 2.98, which indicated the smallest detectable change.

Table 1
Results of the factor analysis.

Item	Somatic symptoms	Affective-cognitive symptoms
1. My appetite was poor.	0.5	0.1
2. I could not shake off the blues.	0.1	0.7
3. I had trouble keeping my mind on what I was doing.	0.3	0.6
4. I felt depressed.	0.2	0.7
5. My sleep was restless.	0.6	0.2
6. I felt sad.	0.1	0.6
7. I could not get going.	0.8	−0.2
8. Nothing made me happy.	0.3	0.4
9. I felt like a bad person.	−0.3	0.6
10. I lost interest in my usual activities.	0.2	0.5
11. I slept much more than usual.	0.5	−0.1
12. I felt like I was moving too slowly.	0.9	−0.3
13. I felt fidgety.	0.4	0.3
14. I wished I were dead.	−0.2	0.5
15. I wanted to hurt myself.	−0.2	0.6
16. I was tired all the time.	0.7	−0.01
17. I did not like myself.	−0.2	0.6
18. I lost a lot of weight without trying to.	0.7	−0.3
19. I had a lot of trouble getting to sleep.	0.5	0.1
20. I could not focus on the important things.	0.3	0.4
% of total variance explained	32.3	9.2

3.8. Responsiveness and floor/ceiling effect

The value of SDC (2.98) was lower than that of MIC (10), which demonstrated the adequate responsiveness of the CESD-R. There was 11.9% (37/310) of total participants that achieved the lowest possible score (0). No participant (0%) achieved the highest possible score on the CESD-R (80). The total percentages of respondents that achieved the lowest or highest possible score were both below 15%, indicating that there were no floor or ceiling effects of the CESD-R.

4. Discussion

In this study, we developed the Chinese version of the CESD-R and examined its validity and reliability by administering the scale on a sample of patients with cancer. Our results indicate sufficient validity (cross-cultural validity, structural validity, and construct validity), satisfactory reliability (internal consistency and test-retest reliability), adequate responsiveness, and no floor/ceiling effect of the CESD-R.

The process of developing the Chinese version of the CESD-R followed the guidelines prescribed in the Translation and Cultural Adaptation-Principles of Good Practice (Wild et al., 2005) and produced excellent cross-cultural validity based on the COSMIN checklist (Terwee et al., 2012). Good construct validity was indicated by confirming both of the hypotheses. First, a strong, positive correlation between the CESD-R and PHQ-9 scores was found. The PHQ-9 is a sufficiently reliable and valid instrument for assessing the nine symptoms contained in the diagnostic criteria of depression (Chen et al., 2010; Kroenke et al., 2001; Sun et al., 2017; Zhang et al., 2016). Secondly, patients who received chemotherapy and radiotherapy reported higher levels of somatic and affective-cognitive symptoms of depression than patients who did not. This is consistent with prior findings that anti-cancer treatments, such as chemotherapy and radiotherapy, can increase somatic symptoms, including fatigue and sleep disturbance, as well as the affective symptoms of depression (Bower et al., 2011; Gray et al., 2014; Torres et al., 2013).

The results obtained from the factor analyses suggest a two-factor structure of the CESD-R: namely, somatic symptoms and affective-cognitive symptoms of depression. To some extent, this is consistent with the factor structure of the CESD in patients with head and neck cancer (Chhabria and Carnaby, 2017). In that study, three factors were suggested: depressed affect, somatic/retarded activity, and positive affect. The differences mainly occurred due to the deletion of the positive items in the CESD-R from the CESD (Eaton et al., 2004). Furthermore, the themes of the factors are consistent with previous findings that patients with cancer are characterized by significant somatic symptoms compared with depressed patients without a somatic disease (Nikendei et al., 2017). The two-factor structure is also in line with the results of research into the mechanism of depression that showed that there were some differences in the mechanism between somatic and affective-cognitive symptoms (Norcini Pala et al., 2016). In the general population, a two-factor structure of the CESD-R has also been suggested (Van Dam and Earleywine, 2011), but there were a few differences in the factor contents. The two factors in the general population were the factor of affective symptoms and the factor of somatic-cognitive symptoms of depression. The differences in the factor contents may be because somatic symptoms, such as low appetite, sleep disturbances, and fatigue, are more common in patients with cancer than in the general population,

Table 2
Comparison of the total and factor scores of the CESD-R as related to chemotherapy and radiotherapy (Median (25^{th} quartiles, 75^{th} quartiles)).

	CESD-R	*p*	Somatic factor	*p*	Affective-cognitive factor	*p*
Chemotherapy		**0.002**		**0.003**		**0.007**
Yes (*n* = 166)	10 (4, 19)		7 (3, 12)		3 (0, 8)	
No (*n* = 144)	6 (2, 14)		4 (1, 10)		1 (0, 6)	
Radiotherapy		**0.035**		**0.049**		**0.036**
Yes (*n* = 50)	12 (5, 23)		8 (3, 14)		5 (0, 11)	
No (*n* = 260)	8 (2, 17)		5 (2, 11)		2 (0, 7)	

The Mann-Whitney *U* test was used. CESD-R = the Center for Epidemiologic Studies Depression Scale-Revised.

making their loadings strong enough to appear as an independent factor (Nikendei et al., 2017).

The reliability was sufficient, as indicated by the excellent internal consistency and good test-retest reliability. The Cronbach's alphas were 0.82–0.88, which is far higher than the acceptable coefficient of 0.70 (Terwee et al., 2007). In addition, the value of ICCs of total and subscale scores (0.73−0.81) were beyond the preferable value of 0.7 (Mokkink et al., 2010a, 2010b). The responsiveness of the CESD-R was sufficient as supported by both the smaller value of SDC than MIC and the absence of a floor/ceiling effect.

5. Conclusions

Our results demonstrated good psychometric properties of the Chinese version of the CESD-R among patients with cancer. However, there were also several limitations. First, the sample size of the patients examined for the test-retest reliability in the follow-up was relatively small (n = 33). However, the sample size was sufficient for ICC of 0.81 with power of 90% and two-sided alpha of 0.05 (number of subjects = 3–28) (Bujanga and Baharum, 2017). Second, the responsiveness was tested based on the MIC of CESD not the MIC of the CESD-R. MIC is an important psychometric properties of a measurement. It would allow us to detect a clinically important change of the construct (e.g. depression), not only the statistical significance. An anchor variable should be included in future study to calculate the MIC using the commended predictive modeling approach (Terluin et al., 2015).

The Chinese version of the CESD-R was demonstrated to be a sufficiently valid and reliable tool for assessing depressive symptoms in patients with cancer. It provides nurses and clinicians of oncology with an easily administered tool for screening Chinese-speaking patients with cancer for depressive symptoms. It also provides researchers with an instrument for assessing the severity of depressive symptoms and comparing the symptoms to other populations worldwide.

Conflicts of interest

The authors have no conflicts of interest to disclose.

Acknowledgements

This work was supported by the Natural Science Foundation of Shandong Province [grant number ZR2017MC070]. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The authors thank Dr. Naixue Cui for her support in the form of reviewing the translation of the scale.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijnurstu.2019.04.008.

References

APA, 1994. Diagnostic and Statistical Manual of Mental Disorders, 4th edition American Psychiatric Publishing, Washington, DC.

APA, 2013. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub, Washington DC.

Bao, Y., Li, L., Guan, Y., Wang, W., Liu, Y., Wang, P., Huang, X., Tao, S., Wang, Y., 2017. Prevalence and associated positive psychological variables of anxiety and depression among patients with central nervous system tumors in China: a cross-sectional study. Psychooncology 26 (2), 262–269. https://www.ncbi.nlm.nih.gov/pubmed/27072749.

Berry, D.L., Blonquist, T.M., Hong, F., Halpenny, B., Partridge, A.H., 2015. Self-reported adherence to oral cancer therapy: relationships with symptom distress, depression, and personal characteristics. Patient Prefer. Adherence 9, 1587–1592. https://www.ncbi.nlm.nih.gov/pubmed/26604712.

Bower, J.E., Ganz, P.A., Irwin, M.R., Kwan, L., Breen, E.C., Cole, S.W., 2011. Inflammation and behavioral symptoms after breast cancer treatment: do fatigue, depression, and sleep disturbance share a common underlying mechanism? J. Clin. Oncol. 29 (26), 3517–3522. https://www.ncbi.nlm.nih.gov/pubmed/21825266.

Bujanga, M.A., Baharum, N., 2017. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Arch. Orofacial Sci. 12 (1), 1–11. http://www.dental.usm.my/aos/docs/Vol_12/aos-article-0246.pdf.

Busch, A.M., Wagener, T.L., Gregor, K.L., Ring, K.T., Borrelli, B., 2011. Utilizing reliable and clinically significant change criteria to assess for the development of depression during smoking cessation treatment: the importance of tracking idiographic change. Addict. Behav. 36 (12), 1228–1232. https://www.ncbi.nlm.nih.gov/pubmed/21872401.

Chen, S., Chiu, H., Xu, B., Ma, Y., Jin, T., Wu, M., Conwell, Y., 2010. Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care. Int. J. Geriatr. Psychiatry 25 (11), 1127–1133. https://www.ncbi.nlm.nih.gov/pubmed/20029795.

Chhabria, K.S., Carnaby, G.D., 2017. Psychometric validation of the Center for Epidemiological Studies Depression Scale in Head and Neck Cancer patients. Oral Oncol. 75, 158–162. https://www.ncbi.nlm.nih.gov/pubmed/29224813.

Costello, A.B., Osborne, J.W., 2005. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract. Assess. Res. Eval. 10 (7), 1–9. https://www.pareonline.net/pdf/v10n7.pdf.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16 (3), 297–334. doi:http://dx.doi.org/10.1007/BF02310555.

Eaton, W.W., Smith, C., Ybarra, M., Muntaner, C., Tien, A., 2004. Center for Epidemiologic Studies Depression Scale: review and revision (CESD and CESD-R)., 3rd ed. The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Instruments for Adults, Vol. 3. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 363–377.

Eaton, W.W., Ybarra, M., Schwab, J., 2012. The CESD-R is available on the web. Psychiatry Res. 196 (1), 161. https://www.ncbi.nlm.nih.gov/pubmed/22349650.

Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., Bray, F., 2013. GLOBOCAN 2012 v1.0, Cancer incidence and mortality worldwide: IARC CancerBase No. 11 [Internet]. International Agency for Research on Cancer, Lyon, France. . Accessed August 22, 2018. Available from: http://globocan.iarc.fr/.

Gray, N.M., Hall, S.J., Browne, S., Johnston, M., Lee, A.J., Macleod, U., Mitchell, E.D., Samuel, L., Campbell, N.C., 2014. Predictors of anxiety and depression in people with colorectal cancer. Support. Care Cancer 22 (2), 307–314. https://www.ncbi.nlm.nih.gov/pubmed/24077745.

Hong, J.S., Tian, J., 2014. Prevalence of anxiety and depression and their risk factors in Chinese cancer patients. Support. Care Cancer 22 (2), 453–459. http://www.ncbi.nlm.nih.gov/pubmed/24091720.

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30 (2), 179–185. https://link.springer.com/article/10.1007%2FBF02289447.

Janssen, J., Koekkoek, P.S., Biessels, G.J., Kappelle, J.L., Rutten, G., Cog-ID study group, 2019. Depressive symptoms and quality of life after screening for cognitive impairment in patients with type 2 diabetes: observations from the Cog-ID cohort study. BMJ Open 9 (1) e024696. https://www.ncbi.nlm.nih.gov/pubmed/30782740.

Koziara, K., 2016. Assessment of depressiveness in population. Psychometric evaluation of the Polish version of the CESD-R. Psychiatr. Pol. 50 (6), 1109–1117. https://www.ncbi.nlm.nih.gov/pubmed/28211550.

Krebber, A.M., Buffart, L.M., Kleijn, G., Riepma, I.C., de Bree, R., Leemans, C.R., Becker, A., Brug, J., van Straten, A., Cuijpers, P., Verdonck-de Leeuw, I.M., 2014. Prevalence of depression in cancer patients: a meta-analysis of diagnostic interviews and self-report instruments. Psychooncology 23 (2), 121–130. http://www.ncbi.nlm.nih.gov/pubmed/24105788.

Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. 16 (9), 606–613. http://www.ncbi.nlm.nih.gov/pubmed/11556941.

Lee, S., Lee, H., Oh, S.-T., Lee, K., Park, J.Y., Choi, W.-J., 2016. PS223. Validation of the Korean version of Center for Epidemiologic Studies Depression Scale-Revised. Int. J. Neuropsychopharmacol. 19 (Suppl 1), 81–82. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5616818/.

Linden, W., Vodermaier, A., Mackenzie, R., Greig, D., 2012. Anxiety and depression after cancer diagnosis: prevalence rates by cancer type, gender, and age. J. Affect. Disord. 141 (2-3), 343–351. https://www.ncbi.nlm.nih.gov/pubmed/22727334.

Liu, C.L., Liu, L., Zhang, Y., Dai, X.Z., Wu, H., 2017. Prevalence and its associated psychological variables of symptoms of depression and anxiety among ovarian cancer patients in China: a cross-sectional study. Health Qual. Life Outcomes 15 (1), 161. https://www.ncbi.nlm.nih.gov/pubmed/28818112.

Lynn, C., Chenneville, T., Bradley-Klug, K., Walsh, A.S.J., Dedrick, R.F., Rodriguez, C.A., 2019. Depression, anxiety, and posttraumatic stress as predictors of immune functioning: differences between youth with behaviorally and perinatally acquired HIV. AIDS Care 1–10. https://www.ncbi.nlm.nih.gov/pubmed/30829044.

Mausbach, B.T., Irwin, S.A., 2017. Depression and healthcare service utilization in patients with cancer. Psychooncology 26 (8), 1133–1139. https://www.ncbi.nlm.nih.gov/pubmed/27102731.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., de Vet, H.C., 2010a. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual. Life Res. 19 (4), 539–549. https://www.ncbi.nlm.nih.gov/pubmed/20169472.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., de Vet, H.C., 2010b. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J. Clin. Epidemiol. 63 (7), 737–745. https://www.ncbi.nlm.nih.gov/pubmed/20494804.

Mols, F., Schoormans, D., de Hingh, I., Oerlemans, S., Husson, O., 2018. Symptoms of anxiety and depression among colorectal cancer survivors from the population-based, longitudinal PROFILES Registry: prevalence, predictors, and impact on quality of life. Cancer 124 (12), 2621–2628. https://www.ncbi.nlm.nih.gov/pubmed/29624635.

Moon, J.R., Huh, J., Song, J., Kang, I.S., Park, S.W., Chang, S.A., Yang, J.H., Jun, T.G., 2017. The Center for Epidemiologic Studies Depression Scale is an adequate screening instrument for depression and anxiety disorder in adults with congential heart disease. Health Qual. Life Outcomes 15 (176), 1–7. https://www.ncbi.nlm.nih.gov/pubmed/28874154.

Muzzatti, B., Mella, S., Bomben, F., Flaiban, C., Gipponi, K., Piccinin, M., Busato, S., Annunziata, M.A., 2018. Intensity and prevalence of depressive states in cancer inpatients: a large sample descriptive study. Eur. J. Cancer Care (Engl.) 27 (1) . https://www.ncbi.nlm.nih.gov/pubmed/27357446.

Nikendei, C., Terhoeven, V., Ehrenthal, J.C., Maatouk, I., Wild, B., Herzog, W., Friederich, H.C., 2017. Depression profile in cancer patients and patients without a chronic somatic diseasePsychooncology . . [Epub ahead of print] https://www.ncbi.nlm.nih.gov/pubmed/28603908.

Norcini Pala, A., Steca, P., Bagrodia, R., Helpman, L., Colangeli, V., Viale, P., Wainberg, M.L., 2016. Subtypes of depressive symptoms and inflammatory biomarkers: an exploratory study on a sample of HIV-positive patients. Brain Behav. Immun. 56, 105–113. https://www.ncbi.nlm.nih.gov/pubmed/26883521.

O'Connor, B.P., 2000. SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. Behav. Res. Methods Instrum. Comput. 32 (3), 396–402. https://www.ncbi.nlm.nih.gov/pubmed/11029811.

Radloff, L.S., 1977. The CES-D Scale: A self-report depression scale for research in the general population. Appl. Psychol. Meas. 1 (3), 385–401. doi:http://dx.doi.org/10.1177/014662167700100306.

Roberts, R.E., Vernon, S.W., 1983. The Center for Epidemiologic Studies Depression Scale: its use in a community sample. Am. J. Psychiatry 140 (1), 41–46. https://www.ncbi.nlm.nih.gov/pubmed/6847983.

Schroevers, M.J., Sanderman, R., van Sonderen, E., Ranchor, A.V., 2000. The evaluation of the Center for Epidemiologic Studies Depression (CES-D) scale: depressed and positive affect in cancer patients and healthy reference subjects. Qual. Life Res. 9 (9), 1015–1029. https://www.ncbi.nlm.nih.gov/pubmed/11332223.

Sullivan, D.R., Forsberg, C.W., Ganzini, L., Au, D.H., Gould, M.K., Provenzale, D., Slatore, C.G., 2016. Longitudinal changes in depression symptoms and survival among patients with lung cancer: a national cohort assessment. J. Clin. Oncol. 34 (33), 3984–3991. https://www.ncbi.nlm.nih.gov/pubmed/27996350.

Sun, X.Y., Li, Y.X., Yu, C.Q., Li, L.M., 2017. Reliability and validity of depression scales of Chinese version: a systematic review. Chin. J. Epidemiol. 38 (1), 110–116. http://www.ncbi.nlm.nih.gov/pubmed/28100388.

Terluin, B., Eekhout, I., Terwee, C.B., de Vet, H.C., 2015. Minimal important change (MIC) based on a predictive modeling approach was more precise than MIC based on ROC analysis. J. Clin. Epidemiol. 68 (12), 1388–1396. https://www.ncbi.nlm.nih.gov/pubmed/25913670.

Terwee, C.B., Bot, S.D., de Boer, M.R., van der Windt, D.A., Knol, D.L., Dekker, J., Bouter, L.M., de Vet, H.C., 2007. Quality criteria were proposed for measurement properties of health status questionnaires. J. Clin. Epidemiol. 60 (1), 34–42. https://www.ncbi.nlm.nih.gov/pubmed/17161752.

Terwee, C.B., Mokkink, L.B., Knol, D.L., Ostelo, R.W., Bouter, L.M., de Vet, H.C., 2012. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual. Life Res. 21 (4), 651–657. https://www.ncbi.nlm.nih.gov/pubmed/21732199.

Torres, M.A., Pace, T.W., Liu, T., Felger, J.C., Mister, D., Doho, G.H., Kohn, J.N., Barsevick, A.M., Long, Q., Miller, A.H., 2013. Predictors of depression in breast cancer patients treated with radiation: role of prior chemotherapy and nuclear factor kappa B. Cancer 119 (11), 1951–1959. http://www.ncbi.nlm.nih.gov/pubmed/23512358.

Van Dam, N.T., Earleywine, M., 2011. Validation of the Center for Epidemiologic Studies Depression Scale–Revised (CESD-R): pragmatic depression assessment in the general population. Psychiatry Res. 186 (1), 128–132. https://www.ncbi.nlm.nih.gov/pubmed/20843557.

Velicer, W.F., 1976. Determining the number of components from the matrix of partial correlations. Psychometrika 41 (3), 321–327. https://link.springer.com/article/10.1007%2FBF02293557.

Velicer, W.F., Eaton, C.A., Fava, J.L., 2000. Construct explication through factor or component analysis: a review and evaluation of alternative procedures for determining the number of factors or components. In: Goffin, R.D., Helmes, E. (Eds.), Problems and Solutions in Human Assessment: Honoring Douglas N. Jackson at Seventy. Kluwer Academic Publishers, Boston, pp. 41–71.

Wakefield, C.E., Butow, P.N., Aaronson, N.A., Hack, T.F., Hulbert-Williams, N.J., Jacobsen, P.B., International Psycho-Oncology Society Research, C, 2015. Patient-reported depression measures in cancer: a meta-review. Lancet Psychiatry 2 (7), 635–647. https://www.ncbi.nlm.nih.gov/pubmed/26303561.

Weir, J.P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J. Strength Cond. Res. 19 (1), 231–240. https://www.ncbi.nlm.nih.gov/pubmed/15705040.

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., Translation, I.T.Ff., Cultural, A., 2005. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health 8 (2), 94–104. https://www.ncbi.nlm.nih.gov/pubmed/15804318.

Zhang, J., An, S., Wang, L., Rong, W., Wu, F., Feng, L., Wu, J., 2016. Clinic-based depression screening in hepatocellular carcinoma patients using the PHQ-9 depression questionnaire. Chin. J. Behav. Med. Brain Sci. 25 (7), 646–649. doi:http://dx.doi.org/10.3760/cma.j.issn.1674-6554.2016.07.016.