HHS Public Access

Author manuscript

Psychol Assess. Author manuscript; available in PMC 2018 October 01.

Published in final edited form as:

Psychol Assess. 2017 October ; 29(10): 1189–1200. doi:10.1037/pas0000427.

Validity of the Brief Symptom Inventory-18 (BSI-18) for Identifying Depression and Anxiety in Young Adult Cancer Survivors: Comparison with a Structured Clinical Diagnostic Interview

Christopher J. Recklitis, PhD, MPH[1,3], **Jaime E. Blackmon, MA**[1], and **Grace Chang, MD, MPH**[2,3]

[1]Dana-Farber Cancer Institute, Boston, MA

[2]VA Boston Healthcare System, Brockton, MA

[3]Harvard Medical School, Boston MA

Abstract

The Brief Symptom Inventory-18 (BSI-18) is widely used to assess psychological symptoms in cancer survivors, but the validity of conventional BSI-18 cut-off scores in this population has been questioned. This study assessed the accuracy of the BSI-18 for identifying significant anxiety and depression in young adult cancer survivors (YACS), by comparing it to a “gold standard” diagnostic interview measure. Two hundred and fifty young adult cancer survivors (YACS), age 18–40 completed the BSI-18 and the Structured Clinical Interview for DSM-IV (SCID) interview assessing anxiety and depressive disorders. BSI-18 results were compared to SCID criteria using ROC analyses. Forty-four participants (17.7%) met criteria for ≥ 1 SCID diagnoses, and an additional 20 (8.0%) met criteria for clinically significant SCID symptoms without a diagnosis. General concordance between the BSI-18 GSI scale and SCID diagnosis was good (AUC = 0.848), but the two most widely used BSI-18 case-rules failed to identify a majority of survivors with SCID diagnoses, and no alternative BSI-18 cut-off scores met study criteria for clinical screening. Analyses aimed at identifying survivors with significant SCID symptoms or a SCID diagnosis had similar results, as did analyses examining depression and anxiety separately. The BSI-18 shows good overall concordance with a psychiatric interview, but recommended cut-off scores fail to identify a majority of YACS with psychiatric diagnosis. Clinicians should not rely on the BSI-18 alone as a screening measure for YACS. Alternative BSI-18 scoring algorithms optimized for detecting psychiatric symptoms in YACS may be an important step to address this limitation.

Keywords

BSI-18; cancer survivors; screening; psychological distress; validation

Correspondence: Christopher J. Recklitis, 450 Brookline Avenue, SW-320C, Boston, MA 02215; (p) 617-632-3839; Christopher_Recklitis@dfci.harvard.edu.

The Brief Symptom Inventory-18 (BSI-18; (Derogatis, 2001)) is an 18- item self-report checklist measures developed as a brief screen for psychological symptoms in medical patients. The BSI-18 has been widely used in research and clinical applications with a variety of patient populations (e.g., Derogatis & Melisaratos, 1983; Mustanski, Garofalo, Herrick, & Donenberg, 2007; Petkus et al., 2010; Serber et al., 2012), including oncology samples (Bober et al., 2013; Galdon et al., 2008; Kwak et al., 2013b; Merport, Bober, Grose, & Recklitis, 2012; Michel et al., 2010; Michel & Vetsch, 2015; Zabora et al., 2001; Zabora, 2015; Zeltzer et al., 2009). Application of the BSI-18 to oncology patients is supported by its brevity, its coverage of the critical areas of anxiety and depression, and the published manual's inclusion of normative data for an oncology sample.

Within oncology, the BSI-18 has been applied most consistently to studies of young adult cancer survivors (YACS), commonly defined as 18 to 39 or 40 years of age (Bleyer, 2007; Hall et al., 2012; U.S. Department of Health and Human Services, 2006). A recent review (Michel & Vetsch, 2015) found the BSI-18 to be the most widely used measure in screening studies with this age group of cancer survivors, and both the Childhood Cancer Survivors Study (Zeltzer et al., 2009) and the Swiss Childhood Cancer Survivor Study (Michel et al., 2010), two of the largest cohort studies of YACS, use the BSI-18 as their principle measure of psychological adjustment. In addition, a previous study confirmed the factorial validity of the BSI-18 in a large YACS sample (Recklitis et al., 2006), supports its use and interpretation in this population.

Young adult cancer survivors are a population of particular clinical interest because they have been found to be particularly susceptible to psychological adjustment problems after cancer, with several studies demonstrating higher rates of anxiety and depression compared to same-age peers (Hobbie et al., 2000; Kwak et al., 2013a; 2013b; Michel, Rebholz, von der Weid, Bergstraesser, & Kuehni, 2010; Recklitis et al., 2010; Salsman et al., 2014; Zeltzer et al., 2009). Psychological screening has been recommended to address this risk in YACS (Children's Oncology Group, 2013; Clinton-McHarg et al., 2010; Langeveld, Stam, Grootenhuis, & Last, 2002; Recklitis, O'Leary, & Diller, 2003), but there is limited empirical information to guide selection of screening measures for this population. Though the BSI-is widely applied in this group, there is lack of consensus about its utility as a clinical screening measure for YACS (Michel & Vetsch, 2015), and questions about the accuracy of BSI-18 case rules to identify clinically significant symptoms in YACS have been raised (Merport & Recklitis, 2012; Recklitis & Rodriguez, 2007; Zabora et al., 2001).

As with any medical screening process, the goal of mental health screening is to identify high risk individuals who can be evaluated with a more in-depth follow-up assessement and appropriate clinical management (Morrison, 1992). To identify individuals requiring follow-up evaluation of this kind, the BSI-18 manual recommends the following clinical case-rule: respondents who have either 1) t-score score ≥ 63 on the Global Severity Index (GSI) or 2) a t-score ≥ 63 on any two of the three symptom scales (Somatization, Depression, and Anxiety) should be classified as having clinically significant distress (Derogatis, 2001). Although this standard BSI-18 case-rule is prescribed by the test developer, a second case-rule using the GSI scale only, classifies anyone with GSI t-score ≥ 63 as having significant psychological distress; this GSI t-score ≥ 63 case rule has been commonly applied in studies

of cancer patients (Merport & Recklitis, 2012; Michel et al., 2010; Zabora et al., 2001; Zeltzer et al., 2009) and other populations (Hart et al., 2014; Hopp, Anderson, Krumholz, Gruber-Baldini, & Shulman, 2012; Petkus et al., 2010). While both the standard BSI-18 case-rule and the GSI t-score ≥ 63 case rule are widely used for cancer patients and survivors, at least three studies have reported they are not well-suited for cancer survivors and recommended alternative case-rules (Merport & Recklitis, 2012; Recklitis & Rodriguez, 2007; Zabora et al., 2001). One of these studies, (Zabora et al., 2001) compared the BSI-18 to the BSI as a criterion measure and found that a BSI-18 total GSI t-score ≥ 57 case rule was optimal for identifying significant psychological symptoms in a group of pediatric and adult cancer patients. In two later studies, the standard BSI-18 case rule missed more than 50% of adult survivors of childhood cancer identified as significantly distressed on the SCL-90-R (Merport & Recklitis, 2012; Recklitis & Rodriguez, 2007), and a total GSI t-score ≥ 50 was recommended as optimal for identifying survivors with significant psychological distress.

While these studies cast serious doubt on the validity of standard BSI-18 case rules for detecting distress in YACS, they are limited because the criterion measures they used (BSI and SCL-90-R) were checklist measures and not structured diagnostic interviews. This is a significant limitation because these checklist measures themselves have not been extensively validated with oncology samples. Without a “gold standard” psychiatric diagnostic interview as a criterion measure, it is difficult to ascertain with confidence how best to use the BSI-18 to identify clinically significant problems in YACS. To address this need, we aimed to validate the BSI-18 by comparing it to a psychiatric diagnostic interview, the Structured Clinical Interview for the DSM-IV (SCID; First & Gibbon, 2003; First, Spitzer, Gibbon, & Williams, 2002), so that we could evaluate utility of standard BSI-18 case rule, as well as other previously proposed case rules, and determine the most appropriate cut-off score on the BSI-18 in YACS.

Method

Participants

Participants were YACS followed at a single cancer center and recruited to participate in E-Quest, a study of several previously developed psychological screening methods in YACS (Recklitis, Blackmon, & Chang, 2015). To be eligible for E-Quest, survivors had to be 18–40 years of age, English-speaking, three or more years from their first cancer diagnosis, and two or more years since last cancer therapy. Participants were recruited at a scheduled visit to an oncology clinic or at patient education conference at the cancer center. Survivors with severe cognitive or sensory limitations that would interfere with completing the study measures, and those who were acutely ill at the time were not eligible. After obtaining consent, participants completed several psychological screening measures, including the BSI-18 reported on here, before completing an interviewer administered structured diagnostic interview. All measures were completed during a single study visit, and study procedures were approved by the cancer center’s institutional review board.

A total of 349 eligible survivors were approached and 250 (71.6%) enrolled on the study. Participants were 125 males (50.0%) and 125 females (50.0%) aged 18–40 (mean =

29.45yrs, SD = 9.77; Table 1). Participants' first cancer diagnoses were classified as follows: Hodgkin's lymphoma (20%), leukemia (20%), brain tumor (12.8%), testicular cancer (10.8%), non-Hodgkin's lymphoma (10%), breast cancer (9.6%), sarcoma (8%), Burkitt's lymphoma (2%), kidney cancer or Wilm's tumor (1.6%), pituitary tumor (1.2%), neuroblastoma (0.8%), ovarian cancer (0.8%), gastric/stomach cancer (0.4%), head and neck cancer (0.4%), melanoma (0.4%), retinoblastoma (0.4%), and other (0.8%). Age of first diagnosis ranged from birth to 37 years (mean = 20.51 yrs, SD = 9.77).

Measures

The Brief Symptom Inventory –18 (BSI-18) (Derogatis, 2001)—The BSI-18 is a self-report symptom checklist measure consisting of 18 items taken from the 53-item Brief Symptom Inventory (BSI; Derogatis, 1993), itself a shortened form of 90-item Symptom Checklist-90-Revised (SCL-90; Derogatis, 1994). Each BSI-18 item describes a symptom to be rated by respondents along a five-point scale according to how much they have been bothered by the symptom in the prior week. Scores on the 18 items are summarized on the Global Severity Index (GSI). The BSI-18 also includes three symptom scales: Somatization, Depression, and Anxiety, each comprising six items. Raw scores on the BSI-18 are converted to t-scores based on gender-specific normative data from non-patient community dwelling U.S. adults. The BSI-18 has specific instructions for imputation to account for missing items, but these were not applied in this study as all respondents provided complete BSI-18 data.

To aid in interpretation the BSI-18 manual directs users to regard t-score elevations ≥ 63 (representing the top 9 % of normative respondents) as indicating significant scale elevation on each of the three BSI-18 scales, Depression, Anxiety, Somatization, or on the overall Global Severity Index (GSI). In addition, the BSI-18 provides a clinical case-rule classifying respondents who have a t-score score ≥ 63 on the GSI, or any two of the symptom scales as requiring clinical follow-up (Derogatis, 2001). This case rule was originally developed for the SCL-90-R and subsequently recommended for the BSI and BSI-18, though the manual does not present any empirical support for the case-rule as applied to the BSI-18. Following common practice for the BSI (e.g., Brown, Whiteley, Harper, Nichols, & Nieves, 2015; Endermann, 2005) users of the BSI-18 in oncology (Merport & Recklitis, 2012; Michel et al., 2010; Zabora et al., 2001; Zeltzer et al., 2009) and other populations (Hart et al., 2014; Hopp, Anderson, Krumholz, Gruber-Baldini, & Shulman, 2012; Mustanski, Garofalo, Herrick, & Donenberg, 2007; Petkus et al., 2010) commonly define overall significant symptoms on the BSI-18 using the GSI scale alone, classifying respondent with GSI t-score ≥ 63 as having clinically significant symptoms. Both the published BSI-18 case-rule criteria (t-score score ≥ 63 on the GSI, or any two of symptom scales) and the conventional case-rule of GSI t-score ≥ 63 were evaluated in this study. In addition, two alternative BSI-18 case rules, the GSI t-score ≥ 57 cut-off reported in a study of cancer patients (Zabora et al., 2001) and the GSI t-score ≥ 50 cut-off recommended in two studies of YACS (Merport & Recklitis, 2012; Recklitis & Rodriguez, 2007) and a study of elderly medical patients (Petkus et al., 2010) were also evaluated.

Structured Clinical Interview for the DSM-IV (SCID)—The SCID-I Research Version (SCID-I-RV) is a semi-structured clinical interview for identifying axis-I psychiatric diagnoses based on DSM-IV criteria (American Psychiatric Association, 1994), and has been used as a "gold standard" diagnostic measure in hundreds of published reports (First & Gibbon, 2003; Segal, Hersen, & Van Hasselt, 1994; Williams et al., 1992). The SCID is designed to be administered by a trained interviewer familiar with diagnostic criteria. Items and item sequence closely correspond to clinical decision trees for diagnoses so an individual's responses can be easily mapped onto DSM diagnostic criteria. Following an integrated skip logic, once a respondent denies a critical symptom or otherwise fails to meet a diagnostic criterion, no other items for that diagnosis are administered.

The SCID was administered by a single interviewer, a Master's level clinical research coordinator with several years of experience working with hospitalized psychiatric patients, who was blind to the subjects' responses on the BSI-18. The modular structure of the SCID allows it to be tailored by including only sections of the interview relevant to specific diagnoses. Modules for the most common depression and anxiety diagnoses were included as these have significant overlap with domains of the BSI-18 and they are prevalent symptoms in this population (Boyes et al., 2009; Costanzo et al., 2009; Harrington, Hansen, Moskowitz, Todd, & Feuerstein, 2010; Linden, Vodermaier, MacKenzie, & Greig, 2012; Mitchell, Ferguson, Gill, Paul, & Symonds, 2013). For depressive disorders, the Major Depressive Disorder, Minor Depressive Disorder, and Dysthymia SCID modules were included. For anxiety disorders, SCID modules covering Panic Disorder (with and without Agoraphobia), Agoraphobia (without Panic Disorder), Specific Phobia, Social Phobia, Obsessive Compulsive Disorder, Generalized Anxiety Disorder, and Acute Stress Disorder were included. In addition, we modules for Mixed Anxiety and Depressive Disorder, and Adjustment Disorder were also used as these disorders include both anxiety and depressive symptoms.

The SCID includes an optional screener which asks respondents to answer the initial questions included in several diagnostic modules. The rationale for the screener is to avoid respondents discerning that "yes" answers to initial questions of a module lead to a large number of follow-up questions, and therefore tending to deny initial symptoms as a way to speed up the interview (First, Spitzer, Gibbon, & Williams, 2002). For this study, we adapted the screener to include initial critical item(s) for each module included in the study (Table 2). Screening items for Post-Traumatic Stress Disorder (PTSD) were included in the screener, but the full PTSD module was not administered; this was considered appropriate since the BSI-18 evaluates only current symptoms, not the associated symptom history necessary to link them to a past trauma. Screener items were used as the basis for classifying individuals as having significant psychological symptoms (described below).

Classification of Participants Based on the SCID

SCID interviews were scored using standard algorithms for DSM –IV diagnoses (First et al., 2002). Only current symptoms and diagnoses (present within the past 30 days) were recorded. Based on their SCID responses, participants were classified as falling into one of three categories; 1) Participants with SCID Diagnoses, including participants who met

criteria for at least one SCID diagnosis; 2) Participants with Significant SCID Symptoms, defined as participants who endorsed two or more symptoms on the SCID screener (Table 2), but did not meet criteria for a SCID diagnosis: and 3) Participants with No Significant SCID Symptoms, defined as participants who endorsed fewer than two screener symptoms and did not meet criteria for a SCID diagnosis. Of note, individuals who endorse most of the features of a disorder, are classified on the SCID as having a "sub-threshold diagnosis" if all features are likely to be present, but some cannot be confirmed by the respondent (e.g., respondent cannot recall or respond definitively). Five participants had sub-threshold SCID diagnoses only and were combined with participants who had confirmed diagnoses in all analyses.

Statistical Analysis

For descriptive purposes, sample distribution on the BSI-18 and the SCID diagnostic categories were reported. BSI-18 t-scores were classified as 1) No elevation, (t-scores < 50; 2) Minor elevation, (t-scores between 50 and 56; 3) Moderate elevation, (t-scores between 57 and 62; and 4) Marked elevation, (t-scores ≥ 63). Analyses were performed to describe the classification agreement of the BSI-18 with the SCID, beginning with discrimination between survivors with and without any diagnosis. These analyses were conducted first using the standard BSI-18 case rule from the published manual (GSI t-score or two subscale scores ≥ 63; (Derogatis, 2001) and alternative case rules suggested in previous research. Subsequently, ROC analyses were conducted to evaluate other potential GSI cut-off scores. Similar analyses were also performed separately for anxiety and depressive disorders with ROC curves used to determine the concordance of the BSI-18 Depression scale with SCID depression diagnoses, and the concordance of the BSI-18 Anxiety scale with SCID anxiety diagnoses. After examining the agreement of the BSI-18 with SCID diagnoses, these analyses were repeated in order to evaluate the utility of the BSI-18 for discriminating between survivors classified as having either Significant SCID Symptoms or SCID Diagnosis compared to those in the No Significant SCID symptom group.

For ROC analyses we reported area under the curve (AUC) as an indicator of discrimination of a screening measure across the range of possible scores, with AUC ≥ 0.80 interpreted as reflecting good discrimination and AUC ≥ 0.90 reflecting excellent discrimination (Fan, Upadhye, & Worster, 2006; Streiner & Cairney, 2007). In addition, we reported three conditional probabilities for each cut-off score evaluated: sensitivity, or true positive rate, reflecting a screening test's likelihood of providing a correct positive screening result for individuals who actually have the condition of interest; specificity, or true negative rate, reflecting likelihood a screening test will provide correct negative result individuals who do not have the condition, and Total Percent Correct (sometimes referred to as total predictive value or accuracy) reflecting proportion of correct screening results (Grilo & White, 2011; Metz, 1978; Morrison, 1992; Sox, 1986). Evaluation of these parameters was the focus of the analysis because they, unlike AUC which captures overall concordance between measures, directly reflect the accuracy of screening decisions made using specific cut-off scores (Fan, Upadhye, & Worster, 2006; Morrison, 1992; Streiner & Cairney, 2007).

Selecting an appropriate screening criterion involves a trade-off between prioritizing sensitivity or specificity, and clinical screening programs generally select criteria prioritizing high sensitivity (Katz, Kopek, Waldron, Devins, & Tomlinson, 2004; Murphy et al., 1987) in order to ensure that the largest possible proportion of patients with the condition are correctly classified by the screening instrument. It was determined a priori that a BSI-18 cut-off score or case-rule with minimum sensitivity ≥ 85% and specificity ≥ 75% would be adequate to be recommended for screening purposes. These sensitivity and specificity criteria were selected based on these considerations as well as prior research demonstrating that depression and anxiety checklist measures can achieve high sensitivity while maintaining robust specificity (Katz et al., 2004; Lowe et al., 2003; Williams, Pignone, Ramirez, Perez, & Stellato, 2002). In a review of depression checklist measures, for example, sensitivity of 85% and specificity of 74% was the average reported. Our criteria are also similar to those used in prior studies of psychological screening measures in cancer patients, where sensitivity criteria for example, have been as high as 90% (Hong & Tian, 2013) though more moderate criteria have also been applied (e.g., 80%; Tuunainen et al., 2001).

Results

Distribution of the Sample on the SCID and BSI-18

Of 250 survivors, 44 (17.6%) met criteria for one or more SCID diagnoses, 20 (8.0%) were classified has having Significant SCID symptoms (with no SCID diagnosis), and 186 (74.4%) were classified as having No Significant SCID Symptoms (Table 1). Frequencies of specific diagnoses are provided in the supplemental material to this article. Based on their GSI scores, 163 survivors (65.2%) had no elevation (GSI T-score < 50), 43 survivors (17.2%) had minor elevation (GSI t-score 50 – 56), 20 survivors (8.0%) had moderate elevation (GSI t-score 57 – 62). 24 survivors (9.6%) had marked elevation on the BSI-18 (GSI T-score ≥ 63) which is similar to previous reports using a large epidemiological sample of adult survivors of childhood cancer where the proportion of participants with this level of GSI elevation ranged from 7.4% and 12.3% (Recklitis et al., 2006; Zeltzer et al., 2009). The relationship between BSI-18 scores and the three SCID categories is presented graphically in the supplemental material to this article.

Classification Agreement between the BSI-18 and SCID Diagnoses

Utility of the BSI-18 to identify YACS with one or more SCID diagnoses was first evaluated by determining how accurately previously BSI-18 case rules could identify these survivors. Using these case rules, survivors were classified as “cases” on the BSI-18 and these classifications were compared to the results of the SCID (Table 3). The standard BSI-18 case rule from the published manual (GSI t-score or two of three subscale t-scores ≥ 63) had very poor sensitivity compared to the SCID, identifying only 20 of the 44 survivors with a SCID diagnosis (sensitivity = 0.45). Similarly, the widely used GSI t-score ≥ 63 case-rule had poor sensitivity (0.41). The GSI t-score ≥ 57 case-rule demonstrated only slightly improved detection of SCID diagnoses, with sensitivity of 0.57, and specificity of 0.91. The GSI t-score cut-off ≥ 50 had the most balanced profile of the four published case-rules (sensitivity

= 0.75, specificity = 0.74), but did not demonstrate either sufficient sensitivity or specificity to meet study criteria (sensitivity ≥ 85% and specificity ≥ 75%).

To illustrate how these sensitivities and specificities can translate into actual screening results, we calculated what screening decisions would result if the GSI t-score ≥ 63 cut-off were applied to a hypothetical situation in which 100 survivors (20% with a psychiatric diagnosis) were screened using the BSI-18 (Figure 1). In this example, for the 80 survivors with no psychiatric condition, 78 (97%) with BSI-18 scores < 63 would be classified correctly as "True Negatives." However, we would expect less than half of the survivors who actually have a psychiatric diagnosis, 8 of 20 (41%) to be accurately identified (True Positives). Consequently, the majority of survivors with a psychiatric condition (12 of 20) would be missed by the screening program and sent home without further assessment.

Utility of other possible GSI t-scores cut-off scores for detecting survivors with SCID diagnoses was evaluated with ROC analyses (Figure 2). Overall, the AUC of 0.852 demonstrated good discrimination between participants with and without a SCID diagnosis, but no GSI t-score cut-off met study criteria as a case-rule. (Table 4). Low GSI cut-off scores necessary to accurately identify a majority of survivors with a SCID diagnosis had poor specificity indicating they were "over-diagnosing" individuals who did not have a SCID diagnosis as cases on the BSI-18. For example, a GSI t-score cut-off of ≥ 47 detected 93% (sensitivity) of survivors with a SCID diagnosis, but it only accurately classified 52% (specificity) of those with no SCID diagnosis (Table 4).

ROC analyses were also used to evaluate classification agreement of the BSI-18 Depression scale with SCID depression diagnoses, and the BSI-18 Anxiety scale with anxiety diagnoses. In these analyses AUC values indicated good discrimination between survivors with and without diagnoses (Figure 2). However, for both the BSI-18 Depression and Anxiety scales, the standard GSI t-score ≥ 63 cut-off score failed to identify most survivors with SCID diagnoses, with sensitivity of 0.50 for Depression and 0.30 for Anxiety. To achieve a sensitivity of 85%, cut-off scores needed to be lowered considerably, to ≥ 45 for Depression and ≥ 47 for Anxiety, but at these thresholds specificity was unacceptably low (0.52 for Depression and 0.53 for Anxiety). No potential cut-off score on the Depression or Anxiety scales demonstrated sufficient sensitivity and specificity to meet study criteria. (Complete results of these analyses are provided in the supplemental material to this article).

Agreement of the BSI-18 with Significant SCID Symptoms or SCID Diagnoses

By definition psychiatric disorders are marked by symptoms of sufficient intensity and duration to impair normal functioning, but psychological symptoms can have a negative impact on YACS even when they do not meet a diagnostic threshold (Institute of Medicine, 2007; National Cancer Institute, 2015; Recklitis, Sanchez-Varela, & Bober, 2008; Schultz et al., 2007). For that reason, we evaluated accuracy of the BSI-18 for distinguishing between survivors with no significant symptoms on the SCID compared to those with significant SCID symptoms or a SCID diagnosis. As in the previous analyses, published BSI-18 case rules were found to have very low sensitivity compared to the SCID, with sensitivities ranging from a low of 0.38 (standard BSI-18 case rule: GSI t-score, or two of three subscale t-scores ≥ 63), to a high of 0.69 (GSI t-score cut-off ≥ 50) (Table 3). Utility of other possible

GSI t-scores cut-off scores for detecting survivors with significant SCID symptoms or SCID diagnoses was evaluated with ROC analyses (Figure 2; Table 4). Similar to previous analyses, the GSI scale demonstrated good discrimination (AUC = 0.843; Figure 2), but no GSI t-score met study criteria as a cut-off score (Table 4). Low GSI cut-off scores necessary to accurately detect at least 85% of survivors with significant SCID symptoms or a SCID diagnosis had poor specificity (< .65).

When these ROC analyses were repeated to evaluate accuracy of the BSI-18 Depression scale to identify survivors with significant SCID depressive symptoms or SCID depressive diagnoses and accuracy of the BSI-18 Anxiety scale to identify those with significant anxiety symptoms or anxiety diagnoses, results were similar; the AUC values indicated good discrimination (Figure 2), but the standard GSI t-score cut-off of ≥ 63 failed to identify most survivors with SCID symptoms of depression or anxiety, with sensitivity of .40 for Depression and .31 for Anxiety. Lower cut-off scores with acceptable sensitivity (≥ 85%) had unacceptably low specificities (0.55 for Depression and 0.64 for Anxiety), so that no potential cut-off score on the Depression or Anxiety scales met study criteria. (Complete results of these analyses are provided in the supplemental material to this article).

Evaluating BSI-18 Case-rules using Total Percent Correct Values

Across all analyses, both the standard BSI-18 case-rule (t-score ≥ 63 on GSI or two subscales) and the GSI t-score ≥ 63 case rule had generally favorable total percent correct values compared to other cut-off scores (Table 4). Similarly, the t-score ≥ 63 cut-off score advocated for interpreting BSI-18 Depression and Anxiety scales also had high total percent correct values compared to other cut-off scores (supporting data provided in the supplemental material to this article). While the sensitivities of these case-rules are unacceptably low for clinical screening, our results support their overall accuracy and use in research applications where the goal is to select a case-rule with the fewest incorrect classifications without regard to whether incorrectly classified individuals represent false positives or false negatives. Of note, As the percent correct of any case-rule will vary with the base rate or prevalence of the condition screened for (unlike sensitivity and specificity), the percent correct values reported here will be informative for considering how the BSI-18 will operate only in samples with similar prevalence of psychiatric symptoms and diagnoses.

Discussion

To examine the validity of the BSI-18 as a clinical screening instrument for YACS this study set out compare the BSI-18 to a psychiatric interview. Prior studies comparing the BSI-18 to other symptoms checklist measures have produced inconsistent recommendations, resulting in several different BSI-18 case-rules being promoted as most appropriate for clinical screening in this population. By evaluating these previously published BSI-18 case-rules against a widely accepted diagnostic criterion measure, we anticipated results of this study would guide clinicians interested in applying the BSI-18 to the YACS population by providing empirical support for one or more of the BSI-18 case-rules. The results, however, indicate that none of the previously developed BSI-18 case-rules met study criteria for a clinical screening measure, and that no other suitable BSI-18 case-rule could be found that

met study criteria for identifying YA survivors with psychiatric disorders or significant psychiatric symptoms on the SCID. This was surprising, given the BSI-18 has been widely used in studies of cancer patients and survivors (Bober et al., 2013; Galdon et al., 2008; Michel et al., 2010; Recklitis et al., 2006; Zeltzer et al., 2009), and several prior studies had reported positively on psychometric properties of the BSI-18 and its concordance with other symptom measures (Lancaster, McCrea, & Nelson, 2016; Maruish, 2004).

Differences between our results and those of previous studies are likely due to differences in analytic strategies and criterion measures. Studies examining the relationship of the BSI-18 items to the symptom scales, and the consistency of these relationships across populations (Galdon et al., 2008; Meijer, de Vries, & van Bruggen, 2011; Recklitis et al., 2006; Wang et al., 2010) have validated the BSI-18 factor structure and scales, but without addressing its accuracy as a screening measure. Similarly, reports of positive correlations of the BSI-18 scales with similar scales (Hoffman, Zevon, D'Arrigo, & Cecchini, 2004; Lancaster et al., 2016; Maruish, 2004) reflect on the shared variance of the BSI-18 with other established measures, but they do not specifically assess utility of specific BSI-18 cut-off scores. Moreover, these results can be misleading if the rating scales used as comparators have their own limitations as screening measures (Carlson et al., 2012).

By calculating the sensitivity and specificity of potential BSI-18 case-rules compared to a diagnostic criterions measure, our analyses focused on evaluating the accuracy of the BSI-18 when used for making relevant clinical screening decisions. Results showed that more than half of YACS with a psychiatric disorder in our sample were missed using the two most widely applied BSI-18 case rules. At best, these false negative decisions represent missed opportunities to identify survivors with mental health needs. But their impact may be even more problematic if they falsely reassure users and become a new impediment to an affected survivor obtaining needed mental health care. The fact that the most widely used BSI-18 case rules failed to identify a majority of YACS with psychiatric disorders is very concerning for anyone using the BSI-18 as a screening measure since the purpose of psychological screening is to bring to light symptoms and disorders not identified in standard clinical care (Morrison, 1992). Moreover, our results are similar to those of a study comparing the BSI-18 to the SCID in a community-based geriatric sample (Petkus et al., 2010); they also found overall concordance of the BSI-18 with the SCID was acceptable (AUC ≥ .80) but standard BSI-18 cut-off scores had very low sensitivity for detecting psychiatric disorders. This suggests that this limitation of the BSI-18 for identifying individuals with a psychiatric diagnosis is not particular to our study or the YACS population.

Diagnoses are not the only dimension of mental health important to assess in cancer survivors, but when the BSI-18 misses YACS who meet psychiatric diagnostic criteria, it fails to identify the survivors most likely to need mental health treatment. In addition, the most commonly used BSI-18 case-rules performed similarly when their potential to identify YACS with either a SCID diagnosis or significant psychological symptoms was evaluated. While alternative case rules with more appropriate sensitivity (≥ 85%) were identified, these case-rules had low specificities that would lead to large number of YA survivors without psychiatric diagnosis being referred for unneeded assessment or care. These false positives

are likely to be unmanageable for mental health professionals in a cancer center and could lead to both patient and staff frustration with screening practices.

Our study has several limitation that should be noted. As we have previously described (Recklitis, Blackmon, & Chang, 2015), the E-Quest study included only YACS 18 – 40 years old, so results should not be generalized to other age groups. Similarly, as the sample was drawn from a single cancer center it may not be representative of the larger population. Though the proportion of survivors with marked elevation on the BSI-18 (GSI t-scores ≥ 63) in our sample was similar to reports from a large cohort of adult survivors of childhood cancers (Recklitis et al., Zeltzer et al., 2009), our results may over- or under-represent the prevalence of psychological problems in YACS. However, as sensitivity and specificity of the BSI-18 or any a screening measure are not affected by the prevalence of the condition of interest (Metz, 1978; Morrison, 1992; Sox, 1986), this is not likely to compromise the validity of our results.

We selected the SCID as the criterion measure for the study as is it one of the most reliable and widely used diagnostic measures in psychiatric research. As the SCID requires a highly trained interviewer, it has been widely used in clinical studies but not epidemiological research (Tsuang et al., 2011), though it has been used in validation studies of epidemiological instruments (e.g., Kessler, 2013). The SCID was considered the most appropriate criterion measure for our study focused on validity of the BSI-18, but lack of epidemiological data does mean our SCID results cannot be readily compared to population norms. Future studies assessing both YACS and healthy controls with the same diagnostic measure will be needed to investigate how the prevalence of mental disorders may differ in YACS. To maximize consistency in administration SCID interviews were conducted by a single interviewer, but consequently we are not able to report interrater reliability for the interviews. It should also be noted that the BSI-18 and SCID were administered on the same day, using their standard reference periods—one week and one month respectively. Symptoms of psychiatric disorders are likely to be stable over these relatively short periods of time, but the possibility that changes in symptoms over the course of a month could contribute to the limited agreement observed between the BSI-18 and the SCID measures needs to be considered. Inquiring about symptom timing and duration on all measures and then investigating temporal changes would allow future studies to investigate this possibility. Additionally, it should be noted that when screening measures are compared to a “gold standard” criterion measure as we did here, the measurement error of the criterion measure is typically ignored, and disagreement between measures is attributed to the screening measure alone. This is an oversimplification, as all measures, including the SCID have some degree of error; to address this limitation, future studies comparing the BSI-18 to different diagnostic measures should be encouraged.

Finally, we specified a priori that a BSI-18 case-rule would need to have sensitivity ≥ 85% and specificity ≥ 75% to be acceptable for clinical use, criteria similar to those applied in other studies of psychological screening in cancer patients (Hong & Tian, 2013; Tuunainen et al, 2001). As we have previously argued (Recklitis, Blackmon, & Chang, 2015), these criteria are help to minimize mistakes in screening decisions that negatively impact both individual patients and health care delivery systems. Selection of these criteria are consistent

with the precision typically demonstrated by medical testing, including laboratory tests where sensitivities above 90% are common, and physician diagnoses where an average of 85% of diseased individuals are accurately identified (Basttan et al. 1998; Graber, 2013; Morrison, 1992). In addition there is evidence that many psychological symptom checklist measures do meet or exceed these thresholds (Williams et al., 2002). Nonetheless, we recognize that some readers may prefer to apply different standards for evaluating the BSI-18; by including sensitivity and specificity data for a range of BSI-18 cut-off scores our data should be informative to them in selecting an appropriate cut-off score for their use.

Despite these limitations, the results have important implications for application of the BSI-18 to YACS. Unfortunately, our results indicate previously proposed BSI-18 case-rules fail to detect up to more than half of YACS with psychiatric disorders, and no other potential case-rules had the balance of sensitivity and specificity necessary for clinical screening applications. This lack of consistency between BSI-18 scores and psychiatric diagnoses means that deciding which survivors to refer for mental health follow-up based solely on the BSI-18 will yield too many incorrect decisions, leading us to conclude that the BSI-18 should not be used as a stand-alone screening measure for making clinical decisions regarding YACS.

The BSI-18 was not developed specifically to correspond to DSM diagnoses of depression and anxiety, but its strong association with indicators of psychopathology in our study and many others suggests it could be useful as part of a clinical evaluation of YACS, as long as users are mindful of its imprecise relationship with psychiatric diagnoses. In addition, future research should investigate methods for improving its accuracy as a screening measure. One approach to this should be to examine whether alternative scoring of the BSI-18 items may produce scales better suited for identifying the sub-group of cancer survivors with one or more psychiatric diagnoses. Alternatively, combining the BSI-18 with other measures in a two-step screening process by selecting a low BSI-18 cut-off that has high sensitivity, and adding a second screening measure with high specificity to eliminate false positives may be an important approach to investigate. This second screen could utilize a symptom checklist or interview measure to assess psychological symptoms, or could focus on aspects of psychological function not captured by the BSI-18. In particular, several investigators have argued that adding measures of symptom duration, associated impairment, and expressed need for psychological intervention may increase screening accuracy (Martinez, Andreu, Galdon, & Ibanez, 2015; Mitchell, Baker-Glenn, Granger, & Symonds, 2010; Schaeffeler et al., 2015); these are aspects of adjustment critical to assessing psychiatric diagnoses, and their absence from most symptom checklist measures including the BSI-18 may contribute to lack of agreement with psychiatric nosology.

Finally, results presented here may not reflect on limitations of the BSI-18 measure alone, but also on the difficulty of developing accurate case-finding measures suitable for routine screening of individuals not seeking mental health care. Routine mental health screening for cancer patients and survivors is endorsed or mandated for by a number of individuals and professional groups (American College of Surgeons, 2012; Holland et al., 2010; Nass et al., 2015; Skolarus et al., 2014), but our results and those of similar studies (Hong & Tian, 2013; Mitchell, 2007; Petkus et al., 2010; Recklitis, Blackmon, & Chang, 2015) raise a significant

question regarding the implementation of this kind of "universal screening." That is—to be practical for routine use in a medical setting, screening measures will need to require limited time and other resources, but can brief self-report symptom checklist measures that meet these requirements deliver highly accurate information for making screening decisions? Unless empirical studies demonstrate that these brief screening measures can meet these requirements, it is unclear if mandates for screening oncology patients will clarify or confuse the question of how these patients and their mental health needs are identified.

In a thought-provoking critique, Salmon et al. argue gaps and inconsistencies in the literature on mental health screening in oncology arise because the diagnostic model borrowed from medicine is not appropriate for the assessment of mental health needs and that "…deciding whether a patient has psychological needs and how these should be met, is too complex to be reduced to a simple screen for distress" (Salmon, Clark, McGrath, & Fisher, 2015). While it may be premature to conclude that a diagnostic screening model cannot guide efforts to identify psychological distress in oncology, it is certain that it cannot do so without empirically validated measures. Limitations in the BSI-18 and other screening methods for cancer patients should spur further efforts to identify appropriate measures for this population, but they may also remind us that screening for psychological symptoms in patients not presenting for mental health care is a challenging undertaking with no guarantee it can be accomplished with brief checklist measures most commonly employed for this purpose.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: Funded by the National Cancer Institute (grant 1R21CA161315 to Christopher Recklitis).

References

American College of Surgeons. Cancer Program Standards (Version 1.2.1). 2012 Retrieved from Commision on Cancer: https://www.facs.org/quality-programs/cancer/coc/standards.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Washington, DC: American Psychiatric Association; 1994.

Bleyer A. Young adult oncology: the patients and their survival challenges. CA: A Cancer Journal for Clinicians. 2007; 57(4):242–255. [PubMed: 17626120]

Bober SL, Zhou ES, Chen B, Manley PE, Kenney LB, Recklitis CJ. Sexual function in childhood cancer survivors: a report from Project REACH. Journal of Sexual Medicine. 2013; 10(8):2084–2093. [PubMed: 23679087]

Boyes AW, Girgis A, Zucca AC, Lecathelinais C. Anxiety and depression among long-term survivors of cancer in Australia: results of a population-based survey. Medical Journal of Australia. 2009; 190(7):S94–S98. [PubMed: 19351302]

Brown LK, Whiteley L, Harper GW, Nichols S, Nieves A. Psychological symptoms among 2032 youth living with HIV: A multisite study. AIDS Patient Care STDS. 2015; 29(4):212–219. [PubMed: 25585049]

Carlson LE, Waller A, Mitchell AJ. Screening for distress and unmet needs in patients with cancer: review and recommendations. Journal of Clinical Oncology. 2012

Children's Oncology Group. Long-term follow-up guidelines for survivors of childhood, adolescent and young adult cancers, Version 4.0. Monrovia, CA: Children's Oncology Group; 2013.

Clinton-McHarg T, Carey M, Sanson-Fisher R, Shakeshaft A, Rainbird K. Measuring the psychosocial health of adolescent and young adult (AYA) cancer survivors: a critical review. Health and Quality of Life Outcomes. 2010; 8:25. [PubMed: 20205922]

Costanzo ES, Ryff CD, Singer BH. Psychosocial adjustment among cancer survivors: findings from a national survey of health and well-being. Health Psychology. 2009; 28(2):147–156. [PubMed: 19290706]

Derogatis, LR. BSI, Brief Symptom Inventory : Administration, scoring & procedures manual. Minneapolis, MN: National Computer Systems; 1993.

Derogatis, LR. SCL-90-R : Symptom Checklist-90-R : Administration, scoring & procedures manual. 3. Minneapolis, MN: National Computer Systems, Inc; 1994.

Derogatis, LR. BSI 18, Brief Symptom Inventory 18: Administration, scoring and Procedure Manual. Minneapolis, MN: NCS Pearson, Incorporated; 2001.

Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychological Medicine. 1983; 13(3):595–605. [PubMed: 6622612]

Endermann M. The Brief Symptom Inventory (BSI) as a screening tool for psychological disorders in patients with epilepsy and mild intellectual disabilities in residential care. Epilepsy & Behavior. 2005; 7(1):85–94. [PubMed: 15939672]

Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. Canadian Journal of Emergency Medicine. 2006; 8(01):19–20. [PubMed: 17175625]

First, MB., Gibbon, M. The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). In: Hersen, M., editor. Comprehensive Handbook of Psychological Assessment, 4 Volume Set. Vol. 2. Hoboken, New Jersey: John Wiley & Sons; 2003.

First, MB., Spitzer, RL., Gibbon, M., Williams, JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP). New York: Biometrics Research, New York Psychiatric Institute; 2002.

Galdon MJ, Dura E, Andreu Y, Ferrando M, Murgui S, Perez S, Ibanez E. Psychometric properties of the Brief Symptom Inventory-18 in a Spanish breast cancer sample. Journal of Psychosomatic Research. 2008; 65(6):533–539. [PubMed: 19027441]

Graber ML. The incidence of diagnostic error in medicine. BMJ Quality & Safety. 2013; 22(Suppl 2):ii21–ii27.

Grilo CM, White MA. A controlled evaluation of the distress criterion for binge eating disorder. Journal of Consulting and Clinical Psychology. 2011; 79(4):509. [PubMed: 21707133]

Hall AE, Boyes AW, Bowman J, Walsh RA, James EL, Girgis A. Young adult cancer survivors' psychosocial well-being: a cross-sectional study assessing quality of life, unmet needs, and health behaviors. Supportive Care in Cancer. 2012; 20(6):1333–1341. [PubMed: 21720746]

Harrington CB, Hansen JA, Moskowitz M, Todd BL, Feuerstein M. It's not over when it's over: long-term symptoms in cancer survivors--a systematic review. International Journal of Psychiatry in Medicine. 2010; 40(2):163–181. [PubMed: 20848873]

Hart T, Benn EK, Bagiella E, Arenth P, Dikmen S, Hesdorffer DC, Zafonte R. Early trajectory of psychiatric symptoms after traumatic brain injury: relationship to patient and injury characteristics. Journal of Neurotrauma. 2014; 31(7):610–617. [PubMed: 24237113]

Hobbie WL, Stuber M, Meeske K, Wissler K, Rourke MT, Ruccione K, Kazak AE. Symptoms of posttraumatic stress in young adult survivors of childhood cancer. Journal of Clinical Oncology. 2000; 18(24):4060–4066. [PubMed: 11118467]

Hoffman BM, Zevon MA, D'Arrigo MC, Cecchini TB. Screening for distress in cancer patients: the NCCN rapid-screening measure. Psycho-Oncology. 2004; 13(11):792–799. [PubMed: 15386639]

Hong JS, Tian J. Sensitivity and specificity of the Distress Thermometer in screening for distress in long-term nasopharyngeal cancer survivors. Current Oncology. 2013; 20(6):e570–e576. [PubMed: 24311958]

Hopp J, Anderson K, Krumholz A, Gruber-Baldini A, Shulman L. Psychogenic seizures and psychogenic movement disorders: Are they the same patients? Epilepsy & Behavior. 2012; 25(4): 666–669. [PubMed: 23168090]

Institute of Medicine. Cancer Care for the Whole Patient Meeting Psychosocial Needs. Washington DC: National Academies Press; 2007.

Katz MR, Kopek N, Waldron J, Devins GM, Tomlinson G. Screening for depression in head and neck cancer. Psycho-Oncology. 2004; 13(4):269–280. [PubMed: 15054731]

Kessler RC, Santiago PN, Colpe LJ, Dempsey CL, First MB, Heeringa SG, Ursano RJ. Clinical reappraisal of the Composite International Diagnostic Interview Screening Scales (CIDI-SC) in the Army Study to Assess Risk and Resilience in Service members (Army STARRS). International Jounal of Methods in Psychiatric Researh. 2013; 22(4):303–321.

Kwak M, Zebrack BJ, Meeske KA, Embry L, Aguilar C, Block R, Cole S. Prevalence and predictors of post-traumatic stress symptoms in adolescent and young adult cancer survivors: a 1-year follow-up study. Psycho-Oncology. 2013a; 22(8):1798–1806. [PubMed: 23135830]

Kwak M, Zebrack BJ, Meeske KA, Embry L, Aguilar C, Block R, Cole S. Trajectories of psychological distress in adolescent and young adult patients with cancer: a 1-year longitudinal study. Journal of Clinical Oncology. 2013b; 31(17):2160–2166. [PubMed: 23650425]

Lancaster MA, McCrea MA, Nelson LD. Psychometric properties and normative data for the Brief Symptom Inventory-18 (BSI-18) in high school and collegiate athletes. The Clinical Neuropsychologist. 2016; 30(2):321–333.

Langeveld NE, Stam H, Grootenhuis MA, Last BF. Quality of life in young adult survivors of childhood cancer. Supportive Care in Cancer. 2002; 10(8):579–600. [PubMed: 12436217]

Linden W, Vodermaier A, MacKenzie R, Greig D. Anxiety and depression after cancer diagnosis: Prevalence rates by cancer type, gender, and age. Journal of Affective Disorders. 2012; 141(2): 343–351. [PubMed: 22727334]

Lowe B, Grafe K, Zipfel S, Spitzer RL, Herrmann-Lingen C, Witte S, Herzog W. Detecting panic disorder in medical and psychosomatic outpatients: comparative validation of the Hospital Anxiety and Depression Scale, the Patient Health Questionnaire, a screening question, and physicians' diagnosis. Journal of Psychosomatic Research. 2003; 55(6):515–519. [PubMed: 14642981]

Maruish, ME. The SCL-90-R, the Brief Symptom Inventory (BSI) and the BSI-18. In: Maruish, ME., editor. The use of psychological testing for treatment planning and outcomes assessment. Vol. 3. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2004.

Meijer RR, de Vries RM, van Bruggen V. An evaluation of the Brief Symptom Inventory–18 using item response theory: Which items are most strongly related to psychological distress? Psychological Assessment. 2011; 23(1):193. [PubMed: 21280957]

Merport A, Bober SL, Grose A, Recklitis CJ. Can the distress thermometer (DT) identify significant psychological distress in long-term cancer survivors? A comparison with the Brief Symptom Inventory-18 (BSI-18). Supportive Care in Cancer. 2012; 20(1):195–198. [PubMed: 21928051]

Merport A, Recklitis CJ. Does the Brief Symptom Inventory-18 case rule apply in adult survivors of childhood cancer? Comparison with the Symptom Checklist-90. Journal of Pediatric Psychology. 2012; 37(6):650–659. [PubMed: 22451261]

Metz CE. Basic principles of ROC analysis. Seminars in Nuclear Medicine. 1978; 8(4):283–298. [PubMed: 112681]

Michel G, Rebholz CE, von der Weid NX, Bergstraesser E, Kuehni CE. Psychological distress in adult survivors of childhood cancer: the Swiss Childhood Cancer Survivor study. Journal of Clinical Oncology. 2010; 28(10):1740–1748. [PubMed: 20194864]

Michel G, Vetsch J. Screening for psychological late effects in childhood, adolescent and young adult cancer survivors: a systematic review. Current Opinion in Oncology. 2015; 27(4):297–305. [PubMed: 26049268]

Mitchell AJ. Pooled results from 38 analyses of the accuracy of distress thermometer and other ultra-short methods of detecting cancer-related mood disorders. Journal of Clinical Oncology. 2007; 25(29):4670–4681. [PubMed: 17846453]

Mitchell AJ, Baker-Glenn EA, Granger L, Symonds P. Can the Distress Thermometer be improved by additional mood domains? Part I. Initial validation of the Emotion Thermometers tool. Psycho-Oncology. 2010; 19(2):125–133. [PubMed: 19296462]

Mitchell AJ, Ferguson DW, Gill J, Paul J, Symonds P. Depression and anxiety in long-term cancer survivors compared with spouses and healthy controls: a systematic review and meta-analysis. Lancet Oncology. 2013; 14(8):721–732. [PubMed: 23759376]

Morrison, AS. Screening for chronic disease: Second edition. Lincoln, NB: The University of Nebraska Press; 1992.

Murphy JM, Berwick DM, Weinstein MC, Borus JF, Budman SH, Klerman GL. Performance of screening and diagnostic tests. Application of receiver operating characteristic analysis. Archives of General Psychiatry. 1987; 44(6):550–555. [PubMed: 3579501]

Mustanski B, Garofalo R, Herrick A, Donenberg G. Psychosocial health problems increase risk for HIV among urban young men who have sex wtih men: preliminary evidence of a syndemic in need of attention. Annals of Behavioral Medicine. 2007; 34(1):37–45. [PubMed: 17688395]

Nass SJ, Beaupin LK, Demark-Wahnefried W, Fasciano K, Ganz PA, Hayes-Lattin B, Smith AW. Identifying and addressing the needs of adolescents and young adults with cancer: Summary of an Institute of Medicine workshop. The Oncologist. 2015; 20(2)

National Cancer Institute. Adjustment to Cancer: Anxiety and Distress-Health Professional Version. Coping with Cancer. 2015. Retrieved from http://www.cancer.gov/about-cancer/coping/feelings/anxiety-distress-hp-pdq

Petkus AJ, Gum AM, Small B, Malcarne VL, Stein MB, Wetherell JL. Evaluation of the factor structure and psychometric properties of the brief symptom inventory—18 with homebound older adults. International Journal of Geriatric Psychiatry. 2010; 25(6):578–587. [PubMed: 20013879]

Recklitis CJ, Blackmon JE, Chang G. Screening young adult cancer survivors for distress with the Distress Thermometer: Comparisons with a Structured Clinical diagnostic interview. Cancer. 2015 n/a-n/a.

Recklitis CJ, Diller LR, Li X, Najita J, Robison LL, Zeltzer L. Suicide ideation in adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. Journal of Clinical Oncology. 2010; 28(4):655–661. [PubMed: 19841325]

Recklitis C, O'Leary T, Diller L. Utility of routine psychological screening in the childhood cancer survivor clinic. Journal of Clinical Oncology. 2003; 21(5):787–792. [PubMed: 12610175]

Recklitis CJ, Parsons SK, Shih MC, Mertens A, Robison LL, Zeltzer L. Factor structure of the brief symptom inventory-18 in adult survivors of childhood cancer: results from the childhood cancer survivor study. Psychological Assessment. 2006; 18(1):22–32. [PubMed: 16594809]

Recklitis CJ, Rodriguez P. Screening childhood cancer survivors with the brief symptom inventory-18: classification agreement with the symptom checklist-90-revised. Psycho-Oncology. 2007; 16(5): 429–436. [PubMed: 16929465]

Recklitis CJ, Sanchez-Varela V, Bober S. Addressing psychological challenges after cancer: a guide for clinical practice. Oncology (Williston Park). 2008; 22(11 Suppl Nurse Ed):11–20.

Salmon P, Clark L, McGrath E, Fisher P. Screening for psychological distress in cancer: Renewing the research agenda. Psychooncology. 2015; 24(3):262–268. [PubMed: 25082459]

Salsman JM, Garcia SF, Yanez B, Sanford SD, Snyder MA, Victorson D. Physical, emotional, and social health differences between posttreatment young adults with cancer and matched healthy controls. Cancer. 2014; 120(15):2247–2254. [PubMed: 24888335]

Schaeffeler N, Pfeiffer K, Ringwald J, Brucker S, Wallwiener M, Zipfel S, Teufel M. Assessing the need for psychooncological support: screening instruments in combination with patients' subjective evaluation may define psychooncological pathways. Psycho-Oncology. 2015; 24(12): 1784–1791. [PubMed: 26042392]

Schultz KA, Ness KK, Whitton J, Recklitis C, Zebrack B, Robison LL, Mertens AC. Behavioral and social outcomes in adolescent survivors of childhood cancer: a report from the childhood cancer survivor study. Journal of Clinical Oncology. 2007; 25(24):3649–3656. [PubMed: 17704415]

Segal DL, Hersen M, Van Hasselt VB. Reliability of the Structured Clinical Interview for DSM-III-R: an evaluative review. Comprehensive Psychiatry. 1994; 35(4):316–327. [PubMed: 7956189]

Serber ER, Edwards-Hamton SA, Yeager B, Clair M, Taylor M, Galloway SK, Borckardt JJ. Prevalence of chest pain, depression, somatization, anxiety, global distress, and substance use among cardiac and pulmonary rehabilitation patients. Pain and Research Treatment, 2012. 2012

Skolarus TA, Wolf AM, Erb NL, Brooks DD, Rivers BM, Underwood W 3rd, Cowens-Alvarado RL. American Cancer Society prostate cancer survivorship care guidelines. CA: A Cancer Journal for Clinicians. 2014; 64(4):225–249. [PubMed: 24916760]

Sox JHC. Diagnostic Decision: Probability Theory in the Use of Diagnostic Tests An Introduction to Critical Study of the Literature. Annals of Internal Medicine. 1986; 104(1):60–66. [PubMed: 3079637]

Stein KD, Syrjala KL, Andrykowski MA. Physical and psychological long-term and late effects of cancer. Cancer. 2008; 112(11 Suppl):2577–2592. [PubMed: 18428205]

Streiner DL, Cairney J. What's under the ROC? An introduction to receiver operating characteristics curves. Canadian Journal of Psychiatry. 2007; 52(2):121–128. [PubMed: 17375868]

Tsuang, MT., Tohen, M., Jones, P. Textbook of psychiatric epidemiology. John Wiley & Sons; 2011.

Tuunainen A, Langer RD, Klauber MR, Kripke DF. Short version of the CES-D (Burnam screen) for depression in reference to the structured psychiatric interview. Psychiatric Research. 2001; 103(2–3):261–270.

U.S. Department of Health and Human Services. Closing the Gap: Research and Care Imperatives for Adolescents and Young Adults with Cancer. Washington, DC: U.S. Department of Health and Human Services; 2006. NIH publication 06-6067

Wang J, Kelly BC, Booth BM, Falck RS, Leukefeld C, Carlson RG. Examining factorial structure and measurement invariance of the Brief Symptom Inventory (BSI)-18 among drug users. Addictive behaviors. 2010; 35(1):23–29. [PubMed: 19733442]

Williams J, Pignone M, Ramirez G, Perez S, Stellato C. Identifying depression in primary care: a literature synthesis of case-finding instruments. General Hospital Psychiatry. 2002; 24(4):225–237. [PubMed: 12100833]

Williams JB, Gibbon M, First MB, Spitzer RL, Davies M, Borus J, et al. The Structured Clinical Interview for DSM-III-R (SCID). II. Multisite test-retest reliability. Archives of General Psychiatry. 1992; 49(8):630–636. [PubMed: 1637253]

Zabora J, BrintzenhofeSzoc K, Jacobsen P, Curbow B, Piantadosi S, Hooker C, Derogatis L. A new psychosocial screening instrument for use with cancer patients. Psychosomatics. 2001; 42(3):241–246. [PubMed: 11351113]

Zabora JR. Christ GH. Distress Screening Guidelines for Oncology Social Workers. Handbook of Oncology Social Work: Psychosocial Care for People with Cancer. 2015:115–119.

Zeltzer LK, Recklitis C, Buchbinder D, Zebrack B, Casillas J, Tsao JC, Krull K. Psychological status in childhood cancer survivors: a report from the Childhood Cancer Survivor Study. Journal of Clinical Oncology. 2009; 27(14):2396–2404. [PubMed: 19255309]

Public Significance Statement

The Brief Symptom Inventory-18 (BSI-18) is a self-report checklist measure that has been widely used to assess psychological symptoms in young adult cancer survivors (YACS). In this study, results of the BSI-18 did not reliably identify YACS diagnosed with a psychiatric disorder using an in-person interview. These results suggest that using the BSI-18 alone does not provide an accurate assessment of psychological functioning in this cancer survivor population.

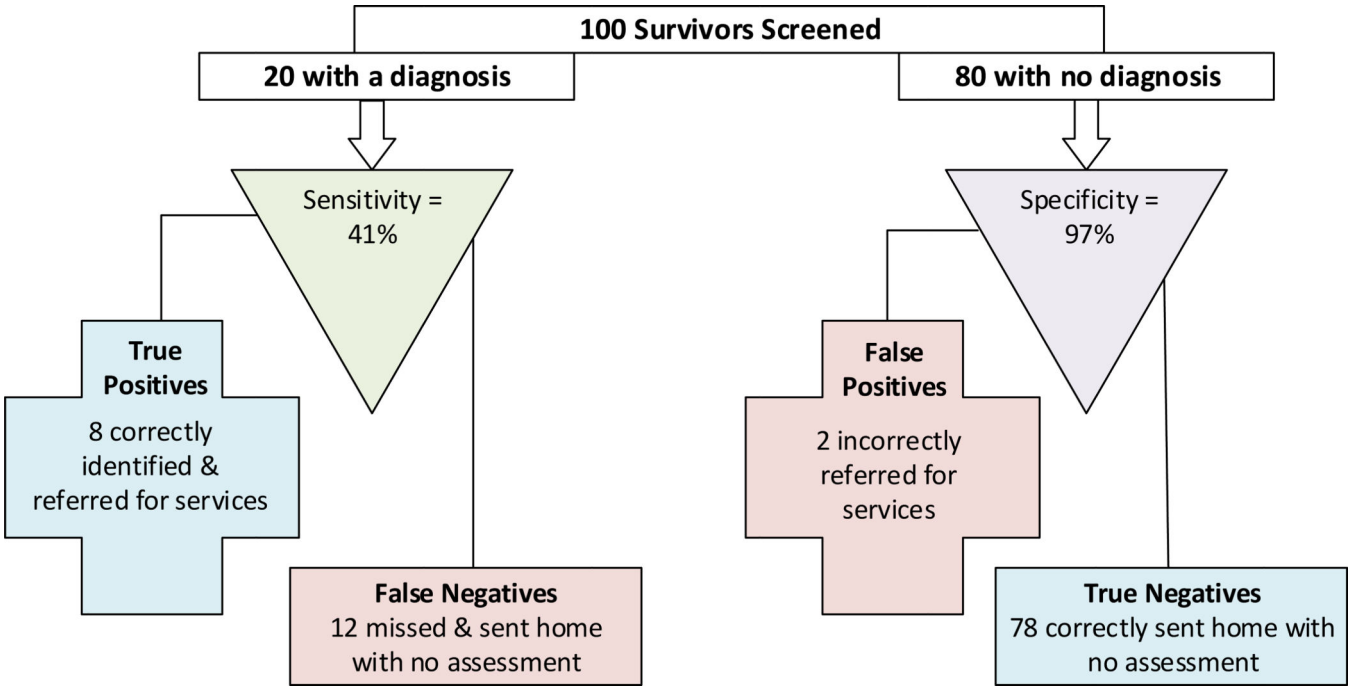

Figure 1.

Expected clinical decisions when using the BSI-18 with a GSI t-score ≥ 63 cut-off to screen young adult cancer survivors. Sensitivity and specificity observed in this sample are applied to a hypothetical example of 100 cancer survivors (20% with a psychiatric diagnosis).

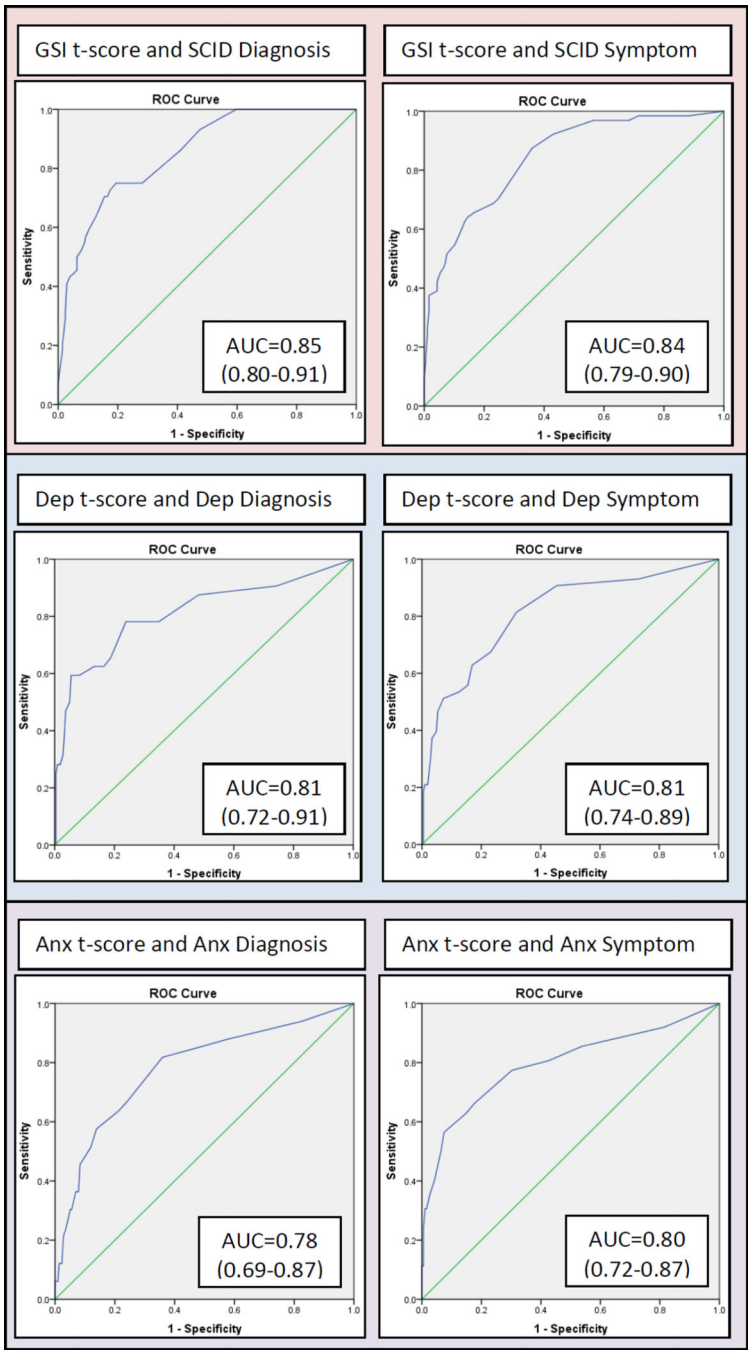

Figure 2.

Receiver operating characteristics (ROC) curve depicting the relationship between the BSI-18 GSI t-score to SCID Diagnosis (left) and SCID Diagnosis or Significant Symptoms (right).

Note. Anx = anxiety; Dep = depression.

Table 1

Description of the Sample (N=250)

Gender	**N**	**%**
Male	125	50.0
Female	125	50.0
Ethnicity		
Caucasian	210	84.0
Hispanic	14	5.6
Asian/Pacific Islander	7	2.8
African-American	6	2.4
Other	10	4.0
Missing	*3*	*1.2*
Age at enrollment		
18–21	55	22.0
22–26	45	18.0
27–31	42	16.8
32–36	41	16.4
37–40	67	26.8
Age at diagnosis		
0–5	20	8.0
6–11	23	9.2
12–17	68	27.2
18–23	37	14.8
24–30	51	20.4
31+	51	20.4
Years since treatment		
2–4	102	40.8
5–9	91	36.4
10–15	27	10.8
15+	28	11.2
missing	2	.8
Cancer diagnosis		
Hodgkin’s Lymphoma	50	20.0
Leukemia	50	20.0
Brain tumor	32	12.8
Non-Hodgkin’s Lymphoma	30	12.0
Testicular	27	10.8
Breast	24	9.6
Sarcomas	20	8.0
Other	17	6.8
	N	**%**
BSI-18 GSI t-scores		

Author Manuscript

Author Manuscript

Author Manuscript

GSI t-score < 50	163	65.2
GSI t-score 50 – 56	43	17.2
GSI t-score 57 – 62	20	8.0
GSI t-score ≥ 63	24	9.6
SCID Criteria	**N**	**%**
SCID Diagnosis	44	17.6
Significant SCID Symptoms no Diagnosis	20	8.0
No Significant SCID Symptoms	186	74.4

Table 2

Critical Symptoms Evaluated on the SCID Screener

Symptom Description	SCID Item	DSM-IV-TR Criteria
Depressive Symptoms		
1. Depressed mood most of the time for 2 weeks in the past month	A1	Major and Minor Depressive Episode Criteria A
2. Diminished interest most of the time for 2 weeks in the past month	A2	Major and Minor Depressive Episode Criteria A
3. Depressed mood most of the time for past 2 years	A163	Dysthymic Disorder Criteria A
Anxiety Symptoms		
4. Recurrent unexpected panic attacks with associated worry or avoidance	F1	Panic Disorder Criteria A
5. Anxiety about being in inescapable places in event of having panic symptoms & agoraphobic symptoms	F29	Agoraphobia Disorder Criteria A
Anxiety Symptoms	**SCID Item**	**DSM-IV-TR Criteria**
6. Marked or persistent fear of social or performance situations	F47	Social Phobia Criteria A
7. Marked or persistent fear of a specific object or situation	F67 and F80	Specific Phobia Criteria A
8. Recurrent, persistent, intrusive, inappropriate thoughts provoking anxiety/distress	F85	OCD Criteria A1
9. Repetitive behaviors the respondent is driven to perform	F89	OCD Criteria A1
10. Excessive anxiety and worry about a number of events/activities > 50% of the time, for ≥ 6 months	S10 and F135	GAD Criteria A
11. Intense fear, helplessness or horror following exposure to a recent traumatic event (past month)	J1 and J2	Acute Stress Disorder Criteria
Anxiety Symptoms	**SCID Item**	**DSM-IV-TR Criteria**
12. Persistent re-experience (nightmares, flashbacks, thoughts) of a traumatic event occurring ≥ 1 month prior	Adapted from F107–109	PTSD Criteria B
13. Intense psychological distress at exposure to cues that resemble a traumatic event occurring > 1 month prior	F110	PTSD Criteria B
Mixed Anxiety/Depressive Symptoms		
14. Persistent or recurrent dysphoric mood state lasting ≥ 1 month	J20	MAD Criteria A

Note. The SCID does not include screening items for Minor Depressive Disorder or Adjustment Disorders. Following standard SCID administration practices, participants who endorsed screener items listed above, but did not meet diagnostic criteria for any disorder, were administered these modules as indicated.

Table 3

BSI-18 Case Rules Compared with SCID Diagnosis and Symptom Cases

	SCID diagnoses			SCID significant symptoms		
	Case (44)	**Non-Case (206)**	**Total (250)**	**Case (64)**	**Non-Case (186)**	**Total (250)**
BSI-18 case rule [a]						
Case	20	7	27	24	3	27
Non-case	24	199	223	40	183	223
GSI t-score ≥ 63 [b]						
Case	18	6	24	21	3	24
Non-case	26	200	226	43	183	226
GSI t-score ≥ 57 [c]						
Case	25	19	44	31	13	44
Non-case	19	187	206	33	173	206
GSI t-score ≥ 50 [d]						
Case	33	54	87	44	43	87
Non-case	11	152	163	20	143	163

[a] **BSI-18 Case Rule:** *Diagnoses, s*ensitivity (SE): 0.45 (0.13–0.23); specificity (SP): 0.97 (0.92–0.99); percent correct (PC): 87.6; *Significant symptoms*, SE: 0.38 (0.26–0.51); SP: 0.98 (0.95–1.00); PC: 82.8.

[b] **GSI t-score ≥ 63:** *Diagnoses*, SE: 0.41 (0.27–0.57); SP: 0.97 (0.93–0.99); PC: 87.2; *Significant symptoms*, SE: 0.33 (0.22–0.46); SP: 0.98 (0.95–1.00); PC: 81.6.

[c] **GSI t-score ≥ 57:** *Diagnoses*, SE: 0.57 (0.41–0.71); SP: 0.91 (0.86–0.94); PC: 84.8; *Significant symptoms*, SE: 0.48 (0.36–0.61); SP: 0.93 (0.88–0.96); PC: 81.6.

[d] **GSI t-score ≥ 50:** *Diagnoses*, SE: 0.75 (0.59–0.86); SP: 0.74 (0.67–0.80); PC: 74.0; *Significant symptoms*, SE: 0.69 (0.56–0.79); SP: 0.77 (0.70–0.83); PC: 74.8.

Table 4

Sensitivity, Specificity, and Percentage Correct of the BSI-18 for Detecting Survivors With A) Any SCID Diagnosis and B) Any SCID Diagnosis or Significant Symptom

BSI-18 GSI t- score cut-off	A. SCID diagnosis (N=44)			B. Diagnosis or significant symptoms (N=64)		
	Sensitivity	Specificity	% Correct	Sensitivity	Specificity	% Correct
≥ 47	.93 (.80–.98)	.52 (.45–.59)	59.6	.92 (.82–.97)	.57 (.50–.64)	66.0
≥ 48	.86 (.72–.94)	.59 (.52–.65)	63.6	.88 (.76–.94)	.64 (.57–.71)	70.0
≥ 49	.75 (.59–.86)	.72 (.65–.78)	72.4	.70 (.57–.81)	.75 (.68–.81)	74.0
≥ 50	.75 (.59–.86)	.74 (.67–.80)	74.0	.69 (.56–.79)	.77 (.70–.83)	74.8
≥ 51	.75 (.59–.86)	.81 (.74–.86)	79.6	.66 (.53–.77)	.83 (.77–.88)	78.8
≥ 52	.73 (.57–.85)	.83 (.76–.87)	80.8	.64 (.51–.75)	.85 (.79–.90)	80.0
≥ 53	.70 (.55–.83)	.83 (.78–.88)	81.2	.63 (.40–.74)	.87 (.81–.91)	80.4
≥ 54	.70 (.55–.83)	.84 (.79–.89)	82.0	.61 (.48–.73)	.87 (.81–.91)	80.4
≥ 55	.64 (.48–.77)	.87 (.82–.91)	83.2	.55 (.42–.67)	.90 (.84–.94)	80.8
≥ 57	.57 (.41–.71)	.91 (.86–.94)	84.8	.48 (.36–.61)	.93 (.88–.96)	81.6
≥ 58	.55 (.40–.69)	.91 (.86–.95)	84.8	.47 (.34–.60)	.94 (.89–.96)	81.6
≥ 59	.52 (.37–.67)	.92 (.87–.95)	85.2	.45 (.33–.58)	.95 (.90–.97)	82.0
≥ 60	.50 (.35–.65)	.94 (.89–.96)	86.0	.42 (.30–.55)	.96 (.91–.98)	82.0
≥ 63	.41 (.27–.57)	.97 (.93–.99)	87.2	.33 (.22–.46)	.98 (.95–1.00)	81.6
≥ 64	.32 (.19–.48)	.98 (.94–.99)	86.0	.27 (.17–.39)	.99 (.96–1.00)	80.4
≥ 65	.30 (.17–.45)	.98 (.94–.99)	85.6	.25 (.15–.38)	.99 (.96–1.00)	80.0