Psychiatry Research 220 (2014) 679–686

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

The differential diagnostic accuracy of the PTSD Checklist among men versus women in a community sample

Kelly S. Parker-Guilbert, Feea R. Leifker, Lauren M. Sippel, Amy D. Marshall *

Department of Psychology, The Pennsylvania State University, 140 Moore Building, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 17 September 2013
Received in revised form
1 August 2014
Accepted 2 August 2014
Available online 9 August 2014

Keywords:
PCL
Assessment
Gender
Trauma
Recruitment
Screening

ABSTRACT

We evaluated the specific version of the PTSD Checklist (PCL-S) as a screening tool for the recruitment of community-residing men and women with diverse trauma experiences. We administered the PCL-S via telephone in the context of participant recruitment, as well as in a laboratory setting preceding administration of the Clinician Administered PTSD Scale (CAPS), the gold standard PTSD assessment tool. In the laboratory, the PCL-S performed reasonably well for men and women, yielding overall diagnostic efficiency (ODE) values (representing percentage of cases accurately identified) of 0.78 and 0.73, respectively, for our recommended cut-points of 42 for men and 49 for women. In contrast, as a recruitment tool, the PCL-S yielded an acceptable ODE of 0.79 for men at the recommended cut-point of 47, but only an ODE of 0.56 (representing diagnostic efficiency no greater than chance) for women at the recommended cut-point of 50. A recruitment cut-point of 57 for women yields a similarly modest ODE of 0.61, but with substantial cost to sensitivity. These findings suggest that use of the PCL-S to screen for PTSD among potential study participants may lead to gender biased study results, even when separate diagnostic cut-points for men and women are used.

1. Introduction

The development of a brief, effective tool to identify individuals with posttraumatic stress disorder (PTSD) is important in multiple ways. In clinical settings, PTSD screening instruments are commonly used to direct patients at elevated risk of having PTSD to further assessment, followed by the provision of resources and treatment, if indicated. In research settings, PTSD screening instruments are often used to identify potential study participants. In both settings, screening instruments are used to identify individuals with probable PTSD (preferably to be later confirmed using a more accurate, but more resource dependent, diagnostic interview), while also identifying individuals without PTSD to exclude those individuals from follow-up service provision or recruitment into the patient group in research studies. Incorrect identification of individuals with or without PTSD may unnecessarily deplete resources and/or lead to invalid study conclusions.

The PTSD Checklist (PCL; Weathers et al., 1993) is the most widely used PTSD screening instrument. Three versions of the PCL exist, with the specific version (PCL-S) aligning most closely with PTSD diagnostic criteria as it references a specified traumatic event and minimizes reporting of psychological distress that is not trauma-related. The PCL-S has demonstrated diagnostic accuracy in select, somewhat homogenous populations, such as veterans, medical patients, and sexual assault survivors (see McDonald and Calhoun (2010)). Knowing how the PCL-S functions among community-residing men and women unselected for trauma type is important, as such a sample provides more broadly generalizable knowledge.

The existence of gender differences in the prevalence, and potentially the etiology and presentation, of PTSD is becoming a recurring theme in the literature, yet little is known about gender's impact on PTSD screening (Brewin, 2005). Researchers have cautioned that the PCL's psychometric properties may differ across genders (Blanchard et al., 1996) and expressed surprise that little is known about how gender affects the PCL's performance (McDonald and Calhoun, 2010). Few researchers have examined gender differences in the PCL's performance. Freedy et al. (2010) examined gender differences in the PCL-C (civilian version, which does not anchor symptoms to a specific traumatic event) and found its performance was superior among men. King et al. (2013) found the PCL-M (military version, which anchors symptoms to stressful military experiences) predicted PTSD similarly among men and women. In another study, the PCL-M demonstrated slightly better (though statistically nonsignificant) performance among men than women (Yeager et al., 2007). Other PTSD screening instruments have been found to perform better among men

* Corresponding author. Tel.: +1 814 863 1752; fax: +1 814 863 7002.
E-mail address: AmyMarshall@psu.edu (A.D. Marshall).

http://dx.doi.org/10.1016/j.psychres.2014.08.001
0165-1781/

than women (Prins et al., 2003), while additional studies have not revealed significant gender differences (Sheeran and Zimmerman, 2002; Calhoun et al., 2010). The dearth of statistically significant differences may be a function of differential gender differences across military and civilian populations, as well as unequal numbers of men and women within samples. Across most studies that report on this issue, group sizes differ by at least a ratio of 2:1.

The prospect that the PCL-S may be less accurate among women than men is concerning since the lifetime PTSD prevalence rate for women is at least twice that of men (McLean et al., 2011); thus the PCL-S may perform most poorly among the majority of individuals with PTSD. It has been suggested that PTSD manifests differently across the genders, such that it is more frequently characterized by comorbid symptoms of depression and generalized anxiety among women than men (Tolin and Foa, 2006). Additionally, depressive and anxiety disorders are more prevalent among women than men (Kessler, 2003). Therefore, compared to men, women's PCL-S scores may reflect a greater degree of generalized distress, as opposed to PTSD-specific distress (Armour et al., 2011). Indeed, screening instruments are typically less effective when discriminating between those with PTSD and those with similar symptoms (which may more frequently be the case among women) than when discriminating between those with PTSD and healthy controls. However, few studies of PTSD screening instruments provide information on psychiatric symptoms among those without PTSD (McDonald and Calhoun, 2010), and none have examined this issue as potentially playing a role in the gender-based differential diagnostic accuracy of a PTSD screening instrument.

1.1. Methods for evaluating diagnostic accuracy

Various elements comprise an accurate screening instrument. Such an instrument possesses an optimal combination of sensitivity (i.e., ability to detect true cases) and specificity (i.e., ability to detect false cases). As sensitivity or specificity increases, the other decreases. The overall diagnostic efficiency (ODE) represents the percentage of cases accurately identified by the screening instrument. Related, positive predictive value (PPV) indicates the proportion of individuals who obtain a score above the designated cut-point on the PCL-S who truly have PTSD. Likewise, negative predictive value (NPV) indicates the proportion of individuals who obtain a score below the designated cut-point who truly do not have PTSD. PPV and NPV are naturally influenced by the sample's base rate of PTSD.

Optimal PCL-S cut-points vary by sample and purpose. For example, sensitivity may be more important than specificity in primary care settings where PTSD screening alerts health care professionals to possible PTSD-related difficulties, thus indicating a relatively lower cut-point. In contrast, specificity may be more important in research settings where false positives deplete limited resources without adding to the PTSD knowledge base. Similarly, specificity is especially important when the screening instrument is used without validation by a diagnostic interview, such that inclusion of false positives may yield misleading or invalid study conclusions. Thus, in some research settings, a higher cut-point may be optimal.

A PCL-S cut-point of 44 is generally accepted as indicative of probable PTSD. However, recommendation of this cut-point came from analyses of a small sample of primarily female motor vehicle accident or sexual assault survivors (Blanchard et al., 1996). Later studies suggest optimal PCL cut-points as low as 30 (e.g., Walker et al., 2002) and as high as 60 (Keen et al., 2008). The variability in cut-points is likely a function of study samples being selected based on trauma type, the intended use of the instrument (e.g., patient care, research screening), sample base rates of PTSD, and the gender composition of the sample examined.

1.2. Current study

Three goals existed for the current study. First, we aimed to provide a broadly generalizable estimate of the PCL-S's diagnostic accuracy for the purpose of research participant recruitment, something that has not been offered by prior literature. To do so, we examined the diagnostic accuracy of the PCL-S in a gender-balanced, mixed-trauma community sample in the context of study recruitment. To be comparable to prior studies examining the PCL's diagnostic accuracy outside of the context of participant recruitment, we also examined the PCL-S's diagnostic accuracy among the same sample when administered in a laboratory setting concurrent with the Clinician Administered PTSD Scale (CAPS). Second, we aimed to compare the PCL-S's diagnostic accuracy between men and women in a more powerful way than prior studies. To do so, we utilized a sample comprised of equal numbers of men and women who were similar in demographic variables. We recruited such a sample by enrolling heterosexual cohabitating couples. Third, we aimed to explore the potential role of non-PTSD psychiatric symptoms in the diagnostic accuracy of the PCL-S among men and women. To do so, we measured general anxiety and depressive symptoms across men and women.

We hypothesized that the PCL-S would appear to be an effective screening instrument when examined across the total sample; however, diagnostic accuracy would be superior for men. We expected that this hypothesized gender difference would be partly due to more severe symptoms of depression and anxiety among women than men. Throughout, we take the perspective of researchers attempting to effectively recruit a community sample of individuals with PTSD. We aim to provide useful guidelines for researchers in a similar position.

2. Methods

2.1. Participants

The current study was conducted as part of a larger project examining the impact of PTSD on relationship functioning. Community participants were recruited using newspaper and internet advertisements (76%), postcards and/or flyers placed in businesses (20%) and an outpatient mental health clinic (4%). Recruitment efforts targeted married or cohabitating heterosexual couples in which at least one partner experienced a stressful life event and who resided in rural or semi-rural communities.

Couples were excluded from the study if neither partner met criteria for probable PTSD (i.e., PCL-S score above 44; $n=122$ couples), they were no longer interested in the study ($n=8$), partners' combined income exceeded $100,000 per year and/or either partner had more than 6 years post-high school education ($n=3$), or they ended their relationship ($n=1$). Income and education restrictions were used to maintain a sample representative of rural communities.

Participants included 128 individuals from 64 couples. On average, participants were 37.06 (S.D.=12.72) years of age, with an individual monthly income of $1731.00 (S.D.=$1522.00), and 14.3 (S.D.=2.32) years of education. Most (68.7%) were employed and few (17%) were students. Participants identified their race/ethnicity as Caucasian (85.9%), African–American (6.3%), biracial/multiracial (3.9%), or Hispanic/Latino (3.9%). Most (72.7%) couples were married. All but three participants experienced an event that met DSM-IV Criterion A for PTSD. Participants' index traumas varied greatly; the most frequent traumas included motor vehicle accidents, intimate partner violence, and loved ones experiencing life-threatening events or sudden death.

2.2. Procedures

Couples contacted the laboratory, then each partner was screened for probable PTSD using the PCL-S (Weathers et al., 1993) over the telephone, as administered in previous studies (e.g., Blanchard et al., 2002). Participants identified their most stressful life event that currently causes the most distress. Once a Criterion A trauma was identified (including reports of fear, helplessness, or horror), participants verbally completed the PCL-S in reference to that event.

Participants recruited into the study completed a laboratory session an average of 1.55 months (S.D.=1.58) after the telephone screening. During the laboratory session, in addition to procedures not contained in the current report, participants