The New York PTSD risk score for assessment of psychological trauma: Male and female versions

Joseph A. Boscarino a,b,c,*, H. Lester Kirchner a, Stuart N. Hoffman d, Jennifer Sartorius a, Richard E. Adams e, Charles R. Figley f

a Center for Health Research, Geisinger Clinic, 100 North Academy Avenue, Danville, PA 17822-4400, USA
b Department of Medicine & Pediatrics, Mount Sinai School of Medicine, New York, NY, USA
c Department of Psychiatry, Temple University School of Medicine, Philadelphia, PA, USA
d Department of Neurology, Geisinger Clinic, Danville, PA, USA
e Department of Sociology, Kent State University, Kent, OH, USA
f School of Social Work, Tulane University, New Orleans, LA, USA

ARTICLE INFO

Article history:
Received 28 November 2011
Received in revised form
14 April 2012
Accepted 17 April 2012

Keywords:
Posttraumatic stress disorder
Psychological trauma
Diagnostic testing
Patient screening

ABSTRACT

We previously developed a new posttraumatic stress disorder (PTSD) screening instrument—the New York PTSD Risk Score (NYPRS). Since research suggests different PTSD risk factors and outcomes for men and women, in the current study we assessed the suitability of male and female versions of this screening instrument among 3298 adults exposed to traumatic events. Using diagnostic test methods, including receiver operating characteristic (ROC) curve and bootstrap techniques, we examined different prediction domains, including core PTSD symptoms, trauma exposures, sleep disturbances, depression symptoms, and other measures to assess PTSD prediction models for men and women.

While the original NYPRS worked well in predicting PTSD, significant interaction was detected by gender, suggesting that separate models are warranted for men and women. Model comparisons suggested that while the overall results appeared robust, prediction results differed by gender. For example, for women, core PTSD symptoms contributed more to the prediction score than for men. For men, depression symptoms, sleep disturbance, and trauma exposure contributed more to the prediction score. Men also had higher cut-off scores for PTSD compared to women. There were other gender-specific differences as well. The NYPRS is a screener that appears to be effective in predicting PTSD status among at-risk populations. However, consistent with other medical research, this instrument appears to require male and female versions to be the most effective.

1. Introduction

The goal of this study is to identify effective risk assessment instruments for posttraumatic stress disorder (PTSD) screening, including assessing the utility of gender-specific instruments. To meet this objective, we used a study of the World Trade Center disaster (WTCD) in New York City (NYC) (Boscarino et al., 2004, 2011a), together with data collected in a chronic pain and in a trauma study (Boscarino et al., 2011b). A number of brief PTSD screening tools are available, including the Primary Care PTSD Screener (PCPS), the Short Screening Scale for PTSD (SSSP), the Abbreviated PTSD Checklist (APCL), and the Short PTSD Rating Interview (SPRINT), among others (Breslau et al., 1999b; Winston et al., 2003; Brewin, 2005; Lang and Stein, 2005; Norris et al., 2006; Bliese et al., 2008; Calhoun et al., 2010). These instruments are relatively short, have been shown to have reasonable specificity and sensitivity, and are focused on screening for core PTSD symptoms. At this time, the PCPS is the most widely used PTSD screener, as it is currently being utilized among both military and civilian populations (Bliese et al., 2008; Calhoun et al., 2010; van Dam et al., 2010; Freedy et al., 2010).

The briefness of the PCPS (4 PTSD items) and its widespread use are clearly the strengths of this instrument. However, these features are also potential weaknesses. The PCPS is a simple, one-dimensional screener focused on key PTSD symptoms, including re-experiencing, avoidance, and arousal symptoms. The PCPS measure has the potential limitation of misclassifying persons.
whose PTSD symptoms might be expressed somewhat differently, such as among men and among those from different ethnic/racial
groups (Adams and Brescario, 2005; Brescario and Adams, 2009; 
Felmingham and Bryant, 2012). When the New York PTSD Risk
Score (NYPRS) was originally developed, the goal was to create a 
screening instrument that was practical and effective in different 
clinical settings and among different populations (Boscario et al.,
2011b). To achieve this we examined multiple risk factors that went 
behind the one-dimensional PTSD screeners in current use
(Brewin, 2005). This multi-factorial approach was consistent with 
the method recently used by Marx et al. (2008) in a study 
designed to predict PTSD among Vietnam veterans.

Based on previous research, the specific focus of this current
study is to assess the suitability of PTSD diagnostic screening 
scores for men and women. Women are known to have higher 
rates of PTSD than men (Kessler et al., 1995), different trauma 
histories (Bromet et al., 1998; Tolin and Foa, 2006), and different 
PTSD-related risk factors (Yehuda, 1999). At this time, evidence 
related to the causes of these gender differences in PTSD is 
limited. Nevertheless, research suggests that these differences 
cannot be explained by the occurrence of sexual assault, preex-
isting depression or anxiety disorder, or by gender-related report-
ing bias (Breslau, 2009). It has been suggested that gender 
differences in neuroticism, depression, and the effects of stressful 
experiences might partly explain the higher vulnerability of 
women to PTSD (Breslau, 2009), but this is unclear at this time.

To date, one study suggests that the PTSD Checklist used among 
older adults seen in primary care clinics should have a lower cutoff 
for female veterans (Lang et al., 2003), but few PTSD screening 
studies have reported differences by gender (Yeager et al., 2007).
Our hypothesis was that the NYPRS would detect a significant 
difference by gender, given the scale’s multi-dimensional struc-
ture. As suggested, most PTSD screeners are not gender-specific 
(Lang et al., 2003). However, these studies typically focus only on 
core PTSD symptoms (Bliese et al., 2008; Calhoun et al.,
2010; van Dam et al., 2010). There is also indirect evidence that 
gender-specific PTSD screeners would be more effective, given 
the differences noted for men and women as this relates to trauma 
exposures, PTSD risk factors, and responses to traumatic stressors
(Kessler et al., 1995; Bromet et al., 1998; Tolin and Foa, 2006;
Breslau, 2009; Maguen et al., 2012). As discussed below, the NYPRS 
includes core PTSD symptoms, as well as depression symptoms, 
trauma history, sleep disturbance, access to care measures, and 
demographic variables. If differences were found by gender for the 
NYPRS, this would mean different risk scores may be warranted for 
men and women with this instrument. In the current study we 
specifically assess the predictive value of different NYPRS models 
for men and for women.

2. Data and methods

2.1. Conceptual approach

Although level of exposure and trauma-related loss are typically associated 
with the impact of traumatic events (Brewin et al., 2000; Norris et al., 2002), 
there are other mediating factors. Research suggests that increased PTSD vulnerability 
ocurs among those with a history of mental health disorders, child adversity, and a history 
of previous traumas (Kessler et al., 1995; Yehuda, 1999; Breslau et al., 1999a). 
Socioeconomic and racial/ethnic factors are also known to affect these experiences
(Adams and Boscario, 2005; Galea et al., 2008). Research has also identified the key 
role of social support among those exposed to traumatic events (Hoboffi et al., 2009;
Adams and Boscario, 2011). The psychobiological bases of this syndrome have also 
become more apparent (Ursano et al., 2010). Given these different psychogenic 
factors, we anticipate a number of health-related problems to emerge among 
traumatized persons, including sleep disturbances, substance misuse, and alterations in 
functional health status, among others.

Our study was in specific response to the National Institute of Mental Health’s 
(NIMH) request for applications related to conducting research with existing 
datasets to develop new PTSD assessment tools (RFA-MH-09-060). To meet this
request, as discussed elsewhere, the NYPRS project team used a multi-factorial 
approach to guide model building combined with diagnost (i.e., atheoretical) 
examinations of statistical results (Boscario et al., 2011b). Noteworthy is that the 
data used by the NYPRS research team utilized data that contained 3298 persons, 
including 270 individuals identified as PTSD-positive cases from three separate 
studies (Table 1).

2.2. Measurements used in the NYPRS

The NYPRS consists of 5 clinical measures plus 5 demographic measures
(Boscario et al., 2011b). The core PTSD screening measure used in the NYPRS is 
the 4-item Primary Care PTSD Screener (PCPS) (Kimerling et al., 2006; Guimette et al.,
2008; van Dam et al., 2010). As suggested, currently this instrument is 
widely utilized among both military and civilian populations (Blesie et al., 2008;
Calhoun et al., 2010; van Dam et al., 2010). The cut-point we used for PTSD with 
the PCPS was 3 positive items, which is the recommended PCPS score to predict 
PTSD (Calhoun et al., 2010). To assess depression symptoms in the NYPRS we used the 
PHQ-2 scale, a commonly used 2-item depression screener (Whooley et al., 1997).
For lifetime trauma exposure we used a simple count of the number of traumatic events (e.g., combat exposure, sexual assault, major disasters, etc.)
experienced in the person’s lifetime. For sleep disturbance we use a single-item 
measure of sleeping problems experienced in the past year. For access to 
healthcare we used a single-item report of access to a regular doctor or to 
medical care services. Demographic measures were based on self report and 
include age, gender, race, Hispanic ethnicity, and college status. Additional 
information on these measures is presented in Table 1 and in the study 
Appendix, as well as published elsewhere (Boscario et al., 2011b).

2.3. Statistical approach

We used a process of moving candidate variables in and out of existing 
statistical models to allow for the manipulation of specificity and sensitivity (Pepe, 2003).
We used methods designed for diagnostic test development, including sensitivity, 
specificity, receiver operator characteristic (ROC) curves, and bootstrapping (Pepe,
2003). An initial model was developed using variables thought to predict PTSD.
This model was then extended to include other candidate measures. These

<table>
<thead>
<tr>
<th>Study variables</th>
<th>WTCD study (N = 2368) % (n)</th>
<th>Geisinger pain study (N = 705) % (n)</th>
<th>Geisinger trauma study (N = 225) % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCPS</td>
<td>23.4 (553)</td>
<td>13.3 (94)</td>
<td>28.0 (63)</td>
</tr>
<tr>
<td>Current PTSD</td>
<td>7.3 (174)</td>
<td>9.9 (70)</td>
<td>11.8 (26)</td>
</tr>
<tr>
<td>PHQ-2 depression score</td>
<td>No symptoms</td>
<td>52.7 (124)</td>
<td>37.0 (261)</td>
</tr>
<tr>
<td></td>
<td>One symptom</td>
<td>16.1 (382)</td>
<td>17.3 (122)</td>
</tr>
<tr>
<td></td>
<td>Two symptoms</td>
<td>31.2 (738)</td>
<td>45.7 (322)</td>
</tr>
<tr>
<td>Lifetime traumatic event exposure</td>
<td>None</td>
<td>28.0 (664)</td>
<td>214 (151)</td>
</tr>
<tr>
<td></td>
<td>Low (&lt; 2)</td>
<td>23.6 (558)</td>
<td>24.8 (175)</td>
</tr>
<tr>
<td></td>
<td>Moderate (2–3)</td>
<td>28.2 (667)</td>
<td>30.9 (218)</td>
</tr>
<tr>
<td></td>
<td>High (4+)</td>
<td>20.2 (479)</td>
<td>22.8 (161)</td>
</tr>
<tr>
<td>Trouble sleeping</td>
<td>32.6 (772)</td>
<td>15.2 (107)</td>
<td>32.0 (72)</td>
</tr>
<tr>
<td>Access to healthcare</td>
<td>88.0 (2084)</td>
<td>98.4 (694)</td>
<td>95.1 (214)</td>
</tr>
<tr>
<td>Mean age (SD)</td>
<td>43.1 (15.5)</td>
<td>54.5 (13.7)</td>
<td>48.4 (16.9)</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>42.9 (1016)</td>
<td>32.9 (232)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>57.1 (1352)</td>
<td>671 (473)</td>
</tr>
<tr>
<td>Race</td>
<td>White</td>
<td>42.9 (1015)</td>
<td>98.4 (694)</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>25.6 (606)</td>
<td>0.9 (6)</td>
</tr>
<tr>
<td></td>
<td>Hispanic/Latino</td>
<td>23.6 (559)</td>
<td>0.4 (3)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>7.9 (188)</td>
<td>0.3 (2)</td>
</tr>
<tr>
<td>College graduate</td>
<td>44.5 (1035)</td>
<td>19.7 (139)</td>
<td>26.7 (60)</td>
</tr>
</tbody>
</table>

* WTCD—World Trade Center Disaster; PCPS—Primary Care PTSD Screener; PHQ-2—Patient Health Questionnaire, 2-item version.
variables initially included mental health status, substance misuse, stress exposures, social/community resources, and functional status measures, among others (Boscarino et al., 2011b). The goal of this model building was to estimate the area under the ROC curve (AUC), while using the fewest number of parameters and the simplest measures to reduce the administrative burden of the final instrument. The AUC was estimated at each step to quantify the prediction accuracy (Hanley and McNeil, 1982). The sequential addition of variables to the base model was evaluated in terms of increasing the AUC (Pepe, 2003).

A non-parametric approach was used to compare the added effects of other variables after the contribution of the base model (DeLong et al., 1988). The results of the model were then used to construct risk scores for PTSD. The properties of the risk scores were examined in terms of sensitivity, specificity, AUC, and by use of a nomogram (Harrell, 2001). A nomogram is a graphical tool used to represent the model and assign regression weights. These weights are the equivalent of standardized beta coefficients in linear regression and are developed from the final prediction models (Harrell, 2001). One problem in estimating diagnostic ability is in using the same dataset from which the model was derived, which can result in overestimation (Efron, 1986). This problem was addressed by estimating a bias-corrected version that used a 1000-sample bootstrap procedure to yield an accurate estimate of the AUC (Harrell, 2001). Specifically, bootstrap 95% confidence intervals (CIs) for the bias-corrected AUC were calculated for the ROC curves reported for the WTCD development sample. This procedure has been shown to be statistically superior to the method of cross-validation and the use of training and validation datasets (Harrell, 2001). In addition to estimating the AUC, we also used Youden’s index (Pepe, 2003). The Youden Index is a summary measure of the ROC curve, as it provides a criterion for choosing a cutoff value for which both sensitivity and specificity are maximized (DeLong et al., 1988; Fluss et al., 2005).

Our earlier analyses suggested that while our overall predictions were good, the model’s results appeared to be different for men and women (Boscarino et al., 2011b). Based on this preliminary finding and other gender research reviewed, in the current study our prediction model was tested for an interaction effect by gender. This was done by comparing the model with gender interaction effects for all the parameters in the model to those without these interaction terms included (Harrell, 2001). If this difference was statistically significant, gender-specific prediction models would be warranted. Originally we assessed sensitivity and specificity as functions of participants’ core PTSD symptom, psychosocial status, and demographic variables. A model was subsequently developed to create a risk score resulting in logistic regression weights. The final NYPDS was based on a prediction model that included the PTSD, depression, trauma exposure, sleep disturbance, access to care, and several demographic measures (Table 1). If statistical interaction was detected by gender, separate models for men and for women would be recommended. The statistical software used in our current study included, SAS, version 9.2 (SAS Institute Inc., 2010), Stata, version 11.2 (Stata Corporation, 2011), and R (R Development Core Team, 2011). Statistical analysis was performed with SPSS (version 19). A p-value of less than 0.05 was considered statistically significant.

2.4. Study subjects

2.4.1. The World Trade Center disaster (WTCD) study

To study the impact of the WTCD event, using random-digit dialing, baseline diagnostic interviews were conducted among NYC adults (18 and older) by telephone one-year after the attacks. For the baseline survey, 2368 residents completed the interview from October, 2002 through December, 2002. This survey was administered using a computer-assisted telephone interviewing (CATI) system. In this study, PTSD was diagnosed based on the full Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) (American Psychiatric Association, 1994). The PTSD measure used was developed for telephone administration and used in previous mental health surveys (Resnick et al., 1993; Acionno et al., 2000; Galea et al., 2002). To meet criteria for PTSD in this study, the person had to meet the full diagnostic criteria for PTSD (Boscarino and Adams, 2008). The validity of this PTSD diagnostic scale was reported to be good compared to the CID (Kilpatrick et al., 1998). Versions of this scale have been used in mental health surveys involving over 15,000 telephone interviews, including several WTCD surveys (Resnick et al., 1993; Galea et al., 2002; Boscarino et al., 2004). The WTCD study served as both a developmental and a validation study. The survey cooperation rate for this survey was estimated to be approximately 57% (American Association for Public Opinion Research, 2008). Additional information related to this study has been published elsewhere (Boscarino et al., 2011b). The WTCD, pain and trauma studies were all approved by the Geisinger IRB. The Geisinger Clinic IRB serves as the current IRB of record for all three of these studies.

3. Results

The studies used in our analyses, tabulated by the risk score measures and PTSD, are shown in Table 1. As seen, the WTCD study tends to differ from the pain and trauma studies in terms of race/ethnicity and level of education. However, the WTCD, pain, and trauma studies are more comparable with respect to lifetime trauma exposure and the prevalence of current PTSD (Table 1). Chi-square (χ²) and t-tests indicated nevertheless that the differences between the WTCD and the combined pain and trauma studies were statistically significant (p < 0.05). Consequently, in the current paper we present the results stratified by the WTCD and the combined pain and trauma studies.

Consistent with our previous work, the study results are presented by the key predictor variables identified. As seen, the Primary Care PTSD Screener (PCPS) is followed by the PCPS plus core psychosocial risk factors. This is followed by the PCPS, plus psychosocial risk factors, plus demographic variables. To formally assess a gender interaction effect, we compared the model with gender interaction terms included for the parameters in the final model (i.e., sleep × gender, trauma × gender, depression × gender, healthcare × gender, age × gender, race × gender, Hispanic × gender, and education × gender) to a model without these interactions. Since these results were statistically significant (χ² = 27.24, d.f. = 12, p = 0.007), separate gender-specific risk score models were developed and are presented below (Tables 2 and 3). For the current study, the pain and trauma studies are combined, as discussed above.

Noteworthy is that for male gender, the PCPS alone is a good predictor of PTSD (Table 2). For example, for men in the WTCD development sample, the PCPS had a specificity of 84.1% and a sensitivity of 89.8% (area under the receiver operating curve, AUC = 0.87). Among the male pain and trauma patients, the PCPS had a specificity of 91.6% and a sensitivity of 95.5% (AUC = 0.935). In the WTCD sample, adding the psychosocial predictors (i.e., depression, trauma exposure, sleep disturbance, and healthcare access) to the model with the PCPS included resulted in a significant improvement, with a specificity of 87.7% and sensitivity of 91.5% (AUC = 0.947, p < 0.0001) for men. This improvement was also observed for men in the pain and trauma studies after psychosocial
predictors were added to the model containing the PCPS, with a specificity of 93.4% and sensitivity of 95.5% (AUC=0.964, \( p=0.0165 \)). However, for men the addition of demographic variables to the model (i.e., college status, age, race, and ethnicity) was neither statistically significant for the WTCD sample (\( p=0.089 \)) nor the pain and trauma samples combined (\( p=0.214 \)) (Table 2).

As can be seen for female gender (Table 3), the PTSD screener alone (i.e., PCPS) is also a good predictor of PTSD. In the WTCD sample, the PCPS had a specificity of 80.8% and a sensitivity of 95.6% (area under the receiver operating curve, AUC,=0.882). Among the female pain and trauma patients, the results for the PCPS had a specificity of 92.4% and a sensitivity of 94.6% (AUC=0.935). In the WTCD sample, adding the psychosocial predictors to the model with the PCPS resulted in a significant improvement, with a specificity of 85.5% and sensitivity of 93.0% (AUC=0.941, \( p<0.0001 \)) for women. Significant improvement was also observed among women in the pain and trauma studies after psychosocial predictors were added to the model, with a specificity of 93.0% and sensitivity of 90.5% (AUC=0.965, \( p=0.0084 \)). For women, the addition of demographic variables to the model was neither statistically significant for the WTCD sample (\( p=0.414 \)) nor the pain and trauma samples (\( p=0.814 \) (Table 3), similar to what was reported for men (Table 2).

Table 4 presents PTSD risk-score results (i.e., the final regression-derived weights) used to generate the gender classification results shown in Tables 2 and 3, respectively. As seen in Table 4 for men (top), a positive score on the PCPS (i.e., 3 or more positive symptoms) is given a base score of 100 (otherwise=0) and the psychosocial and demographic items are given weights (or scores) relative to this score. This scoring is based on the logistic regression analyses, whereby the \( b \) coefficients in these logistic regression models are converted to standardized weights using a nomogram, as discussed. Table 4 also shows the cut-off score for men for a PTSD classification, based on these weights: 100 for the PCPS used alone and 184 for the PCPS + psychosocial predictors. A cut-off score for the addition of demographic variables is not shown for men, since these variables were not statistically significant when added to the model.

### Table 2

Results for men: WTCD, pain and trauma studies using different prediction models

<table>
<thead>
<tr>
<th>Study and prediction model used</th>
<th>Cut-off score</th>
<th>% Specificity</th>
<th>% Sensitivity</th>
<th>PV+</th>
<th>PV-</th>
<th>AUC</th>
<th>AUC 95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PTSD screen only</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTCD (N=1016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPS</td>
<td>100</td>
<td>84.1</td>
<td>89.8</td>
<td>25.8</td>
<td>99.3</td>
<td>0.870</td>
<td>0.829-0.910</td>
<td>-</td>
</tr>
<tr>
<td>Pain and trauma studies (N=356)</td>
<td>100</td>
<td>91.6</td>
<td>95.5</td>
<td>42.9</td>
<td>99.7</td>
<td>0.935</td>
<td>0.882-0.989</td>
<td>-</td>
</tr>
<tr>
<td><strong>PTSD screen + risk factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTCD (N=1016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPS + risk factors</td>
<td>184</td>
<td>87.7</td>
<td>91.5</td>
<td>31.4</td>
<td>99.4</td>
<td>0.947</td>
<td>0.930-0.965</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Pain and trauma studies (N=356)</td>
<td>184</td>
<td>93.4</td>
<td>95.5</td>
<td>48.9</td>
<td>99.7</td>
<td>0.964</td>
<td>0.936-0.993</td>
<td>0.0165</td>
</tr>
<tr>
<td><strong>PTSD screen + risk factors + demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTCD (N=1016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPS + risk factors + demos</td>
<td>263</td>
<td>92.5</td>
<td>91.5</td>
<td>42.8</td>
<td>99.4</td>
<td>0.954</td>
<td>0.937-0.970</td>
<td>0.0890</td>
</tr>
<tr>
<td>Pain and trauma studies (N=356)</td>
<td>263</td>
<td>95.2</td>
<td>77.3</td>
<td>51.6</td>
<td>98.5</td>
<td>0.960</td>
<td>0.927-0.993</td>
<td>0.2137</td>
</tr>
</tbody>
</table>

* AUC=Area under ROC curve; PV+ =Predictive value of positive test; PV- =Predictive value of negative test; PCPS=Primary care PTSD Scale; WTCD=World Trade Center Disaster; Demos=Demographics.

### Table 3

Results for women: WTCD, pain and trauma studies using different prediction models

<table>
<thead>
<tr>
<th>Study and prediction model used</th>
<th>Cut-off score</th>
<th>% Specificity</th>
<th>% Sensitivity</th>
<th>PV+</th>
<th>PV-</th>
<th>AUC</th>
<th>AUC 95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PTSD screen only</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTCD (N=1352)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPS</td>
<td>100</td>
<td>80.8</td>
<td>95.6</td>
<td>31.6</td>
<td>99.5</td>
<td>0.882</td>
<td>0.860-0.904</td>
<td>-</td>
</tr>
<tr>
<td>Pain and trauma studies (N=574)</td>
<td>100</td>
<td>92.4</td>
<td>94.6</td>
<td>64.8</td>
<td>99.1</td>
<td>0.935</td>
<td>0.902-0.968</td>
<td>-</td>
</tr>
<tr>
<td><strong>PTSD screen + risk factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTCD (N=1352)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPS + Risk Factors</td>
<td>139</td>
<td>85.5</td>
<td>93.0</td>
<td>37.4</td>
<td>99.3</td>
<td>0.941</td>
<td>0.928-0.955</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Pain and trauma studies (N=574)</td>
<td>139</td>
<td>93.0</td>
<td>90.5</td>
<td>65.7</td>
<td>98.5</td>
<td>0.965</td>
<td>0.950-0.980</td>
<td>0.0084</td>
</tr>
<tr>
<td><strong>PTSD screen + risk factors + demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTCD (N=1352)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPS + risk factors + demos</td>
<td>166</td>
<td>87.5</td>
<td>91.2</td>
<td>40.4</td>
<td>99.1</td>
<td>0.943</td>
<td>0.929-0.956</td>
<td>0.4140</td>
</tr>
<tr>
<td>Pain and trauma studies (N=574)</td>
<td>166</td>
<td>94.0</td>
<td>83.8</td>
<td>65.4</td>
<td>97.5</td>
<td>0.965</td>
<td>0.951-0.980</td>
<td>0.8142</td>
</tr>
</tbody>
</table>

* AUC=Area under ROC curve; PV+ =Predictive value of positive test; PV- =Predictive value of negative test; PCPS=Primary care PTSD Scale; WTCD=World Trade Center Disaster; Demos=Demographics.
4. Discussion

We examined different clinical domains, including PTSD symptoms, stressor exposures, depression symptoms, sleep disturbances, and demographic variables to evaluate different PTSD prediction models by gender. As suggested, our overall study goal was to develop a prediction tool that was useful in clinical practice (Boscarino et al., 2011b). As shown in Tables 2 and 3, five prediction domains were identified, including core PTSD symptoms (i.e., the PCPS), sleep disturbance, current healthcare access, depression symptoms, and past trauma exposure. By itself, the PCPS performed relatively well for both men and women, with the area under the ROC curve (AUC) ranging from 0.870 to 0.935.

The interaction detected for the NYPRS by gender dictated that different models should be considered for men and women \((p=0.007)\). For men in the WTCD study, adding healthcare access, sleep disturbance, depression symptoms, and trauma exposure to this model increased the AUC from 0.870 to 0.947, a significant improvement in the prediction results \((p<0.0001)\). Adding demographic variables increased the AUC to 0.954 for men, which was not statistically significant \((p=0.089)\). As discussed, to validate these findings for men, we applied the prediction results not only to the adults in the WTCD study, but also to those in the pain and trauma studies recruited from a large healthcare system. In summary, the results for men were essentially the same in both the WTCD and the combined pain and trauma studies (Table 2).

For women in the WTCD study, adding healthcare access, sleep disturbance, depression symptoms, and trauma exposure to the model with the PCPS increased the AUC from 0.882 to 0.941, also a significant improvement \((p<0.0001)\). Adding demographic variables increased the AUC to 0.943, which was not significant \((p=0.414)\). As with men, to validate these findings for women, we also applied the prediction results to not only the adults in the WTCD study, but to those in the pain and trauma studies. In summary, these results were essentially similar to the WTCD study results (Table 3).

Examination of the specific prediction weights reveals that men received higher weights for depression symptoms, sleeping problems, and trauma exposures, compared to women. Conversely, women received higher relative prediction weights for poorer healthcare access and for PTSD symptoms \((i.e., \text{the PCPS})\) compared to men (Table 4). Also, the PTSD cut-off score for men = 184, while for women the cut-off score = 139. The latter findings is consistent with a report that suggested different cut-points should be used for men and women for the PTSD Checklist (Lang et al., 2003).

The use of PTSD screeners has increased recently with growing interest in the impact of traumatic stressors in primary care.
Currently, the Department of Veterans Affairs and the Department of Defense are routinely using the PCPS in clinical practice to assess veterans and active duty personnel (Calhoun et al., 2010). As seen in the current study for both men and women, the PCPS screener appears to work well with non-veterans and non-active duty personnel. It is noted that ~8% of the adults in the combined WTCD, pain, and trauma studies, however, are US military veterans, but this group is not analyzed separately, due to the small number of female veterans and the current sample design. The addition of psychosocial risk predictors increases the predictive ability of the PCPS, but adding demographic variables did not. The PCPS consists of 4 PTSD symptom questions, which would require only a few minutes to administer in most cases. If the psychosocial questions are added, which include 2 depression questions, a trauma question, a sleep question, and a healthcare access question, this would still likely require fewer than 5 min for administration. The instrument, minus the demographic variables, would consist of 9 questions, achieving the brevity objective for this new screening scale.

Our hypothesis was that the NYPRS would detect a significant difference by gender, given the multi-dimensional structure of this scale and previous research. As discussed, there was indirect evidence that gender-specific PTSD screeners would be more effective, given the differences found for men and women as this relates to trauma exposure, PTSD risk, and response to traumatic stressors (Kessler et al., 1995; Bromet et al., 1998; Tolin and Foa, 2006; Breslau, 2009; Maguen et al., 2012). Our research suggests that male and female versions of the NYPRS may be more beneficial than simply using one-dimensional screening scales, such as the PCPS. As was seen, for both men and women adding key psychosocial screening elements to the NYPRS significantly increased the positive predictive value of this scale for both groups. The disadvantage of this is that it would add several minutes to the patient interview.

The current study has several strengths and limitations. A major strength was that our study involved a large-scale random survey among a multi-ethnic urban population and two validation studies. These validation studies included the WTCD bootstrap-validation and the combined pain/trauma validation studies. We also assessed a range of psychological and interpersonal risk factors using standardized instruments and medical test methods. In addition, the PTSD reference standard used in our studies has been clinically validated (Resnick et al., 1993; Galea et al., 2002; Boscarino et al., 2004).

Potential study limitations include that we omitted individuals without a telephone, and those who were institutionalized, homeless, or too ill to be interviewed. In addition, non-response bias also could have affected all our survey results (Boscarino et al., 2011b). Furthermore, we did not use the predictors identified to predict PTSD beyond the first year after trauma exposure (i.e., the baseline study), which was within the same timeframe that PTSD was assessed. However, we recently investigated use of a “modified” NYPRS to forecast PTSD two years after initial exposure, which appeared to work very well (Boscarino et al., 2012). The sample sizes for the pain and trauma studies were also relatively small and these studies were also quite different from the WTCD study. These factors could have biased the results. There are other limitations that we have noted (Boscarino et al., 2011b).

Despite these limitations, our study suggests that a simple, brief screening instrument, The New York PTSD Risk Score (NYPRS), male and female versions, may be effective in PTSD screening. We developed gender-specific PTSD risk scores based on use of the PCPS, depression symptoms, sleep disturbance, trauma history, and access to healthcare. This screening instrument had good sensitivity and specificity and was effective in discriminating PTSD cases from non-cases. The NYPRS can be used based on the available patient and/or provider time, including the PCPS alone or in combination with psychosocial predictors. Important is the fact that while the PCPS alone appears to perform well in predicting PTSD cases, the addition of several brief psychosocial risk measures (e.g., depression symptoms, trauma history, care access, and sleep disturbance), significantly improves the prediction results for both men and women. Furthermore, the prediction weights appear to be different for men and women. We think the latter may have important clinical significance in the future. For example, research suggests that men and women have a different responses to PTSD cognitive therapy, express PTSD symptoms differently, have significant differences in the amounts and types of trauma experienced, and have different PTSD risk factors (Maguen et al., 2012; Felmingham and Bryant, 2012). These factors may affect treatment interventions.

The goal of our original effort was to develop risk assessment tools that were sensitive to both statistical and clinical significance in order to develop data useful for clinical decision-making. Our objective was to develop PTSD prediction tools to facilitate intervention by making it possible to identify high-risk cases from among all persons exposed to trauma. Our current study suggests that more effective PTSD screening should involve use of different screeners for men and for women, which has rarely been the case in the past. As shown, women have lower cut-off scores for PTSD than men and their PTSD score weighting is different. The latter makes sense, given the past differences reported for men and for women related to the epidemiology and etiology of PTSD and their different treatment responses (Kessler et al., 1995; Bromet et al., 1998; Tolin and Foa, 2006; Felmingham and Bryant, 2012).

Our study suggests that gender-specific PTSD screeners might be more effective for use among at-risk populations. Further research is recommended to verify our findings and to make the appropriate adjustments to the NYPRS to achieve more effective screening and treatment. At the very least the NYPRS informs the clinician that for men being screened for PTSD, depression symptoms and sleep disturbance may be the critical issues to be addressed, for women PTSD symptoms and healthcare access may be the critical issues. Additional research is advised.

Acknowledgments

This research was supported in part by grants from the National Institute of Mental Health (Grant # R01-MH-66403 and R21-MH-086317), Pennsylvania Department of Health (Contract # 4100042573), and the Geisinger Clinic Endowment (Grants # SRC-041 and # TRA-015), Boscarino PI.

Appendix. New York PTSD risk score (gender version)

Primary Care PTSD Screener (PCPS) (3 positive symptoms out of 4, past 12 months)

1. You had repeated bad dreams or nightmares or had disturbing or unpleasant memories, thoughts, or images that kept coming into your mind whether you wanted to think of them or not.
2. You deliberately tried hard not to think about something that happened to you or went out of your way to avoid certain places or activities that might remind you of something that happened in the past.
3. You felt you had to stay on guard much of the time or unexpected noises startled you more than usual.
4. You felt cut off from other people, found it difficult to feel close to other people, or you could not feel things anymore or you had much less emotion than you used to have.

**Depression Symptoms** (lifetime)

1. Have you ever had a period of two weeks or longer when you were feeling depressed or down most of the day or nearly everyday?
2. Have you ever had a period of two weeks or longer when you were uninterested in most things or unable to enjoy things you used to do?

**Trauma Exposure** (lifetime)

1. How many traumatic life events do you think you have ever experienced? These are events outside of everyday experiences and include being in combat or a war zone, being assaulted or sexually attacked, being in a major disaster, fire, or accident, experience the sudden and unexpected death of a loved one, and things like these.

   Would you say you never experienced these events, experienced these events once, you experienced these events twice, experienced these events three times or you experienced these events four times or more in your lifetime?

**Sleep Disturbance** (past 12 months)

1. You had difficulty falling asleep or staying asleep?

**Source of Healthcare/Regular Doctor** (current)

1. Do you have a regular doctor or a usual source of care that you used to do?

**Demographics**

1. What is the highest level of education or schooling you completed (record as college graduate vs. not college graduate)?
2. How old are you (record in years)?
3. Are you of Spanish or Hispanic origin?
4. How would you describe your racial background: White, Black/African American, Asian, or something else (record as White vs. not White)?
5. Patient/person’s gender (record by observation): Female or Male?

**References**


