Received: 27 December 2015 | Accepted: 20 June 2016

DOI 10.1111/appy.12248

WILEY ASIA-PACIFIC PSYCHIATRY

ORIGINAL ARTICLE

Diagnostic accuracy of Japanese posttraumatic stress measures after a complex disaster: The Fukushima Health Management Survey

Yuriko Suzuki[1*] | Hirooki Yabe[2] | Naoko Horikoshi[3] | Seiji Yasumura[3] | Norito Kawakami[4] | Akira Ohtsuru[5] | Hirobumi Mashiko[6] | Masaharu Maeda[7] | on behalf of the Mental Health Group of the Fukushima Health Management Survey[†]

[1] Department of Adult Mental Health, National Institute of Mental Health, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan

[2] Department of Neuropsychiatry, Fukushima Medical University, Fukushima, Fukushima, Japan

[3] Department of Public Health, Fukushima Medical University, Fukushima, Fukushima, Japan

[4] Department of Mental Health, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

[5] Department of Radiation Health Management, Fukushima Medical University, Fukushima, Fukushima, Japan

[6] Fukushima Prefecture Developmental Disability Support Center, Koriyama, Fukushima, Japan

[7] Department of Disaster Psychiatry, Fukushima Medical University, Fukushima, Fukushima, Japan

Correspondence
Yuriko Suzuki, Department of Adult Mental Health, National Institute of Mental Health, NCNP 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan.
Email: yrsuzuki@ncnp.go.jp

Abstract

Background The Posttraumatic Stress Disorder (PTSD) Checklist (PCL) has been widely used among traumatized populations to screen people with PTSD; however, the Japanese version of the PCL has yet to be validated. We examined the diagnostic accuracy of the Japanese version PCL-Specific (PCL-S) and the abbreviated versions of the PCL-S among the evacuees of the Fukushima Daiichi Nuclear Power Plant accident.

Methods Fifty-one participants were recruited from an evacuee and clinical sample. The PCL-S, Impact of Event Scale-Revised (IES-R), and World Health Organization Composite International Diagnostic Interview were administered. Screening properties of the PCL-S, IES-R, and abbreviated PCL-S against PTSD diagnosis, including sensitivity, specificity, and diagnostic efficiency, were calculated. Receiver operating characteristic curves were drawn, and optimal cutoff points were examined.

Results The sensitivity, specificity, and diagnostic efficiency of the PCL-S were 66.7%, 84.9%, and 79.2%, respectively (at 52, the area under the curve was 0.83). The cutoff point method for the PCL-S performed better than did the symptom cluster method. The screening properties of the abbreviated versions were comparable with those of the full version.

Conclusions The Japanese version of the PCL-S showed moderate diagnostic accuracy and improved performance over the IES-R for PTSD diagnosis based on the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition. The Japanese version of the PCL-S was a reliable and valid measure, and its diagnostic accuracy was reasonable for both full and abbreviated versions.

KEYWORDS

Fukushima Nuclear Accident, mass screening, posttraumatic stress disorders, receiver operating characteristic curve, sensitivity and specificity

In the aftermath of a disaster, posttraumatic stress disorder (PTSD) is a common and important psychiatric disorder (North & Pfefferbaum, 2013). The prevalence of probable PTSD has been estimated in the range of 2.3% to 44.6%, depending on the population, type of trauma, elapsed period since trauma exposure, and instrument used (Neria, Nandi, & Galea, 2008). In estimating the prevalence of PTSD, various traumatic stress instruments have been used, including the PTSD Checklist (PCL) (Weathers, Litz, Herman, Hushka, & Keane, 1993) and Impact of Event Scale-Revised (IES-R) (Weiss & Marmar, 1997). The diagnostic accuracy of these instruments differs according to the characteristics of the target population and the base rate of PTSD (Terhakopian, Sinaii, Engel, Schnurr, & Hoge, 2008). Thus, it is important to calibrate the instrument and examine the optimal cutoff point, depending on the study population and context.

[†] Membership of the Mental Health Group of the Fukushima Health Management Survey is provided in the acknowledgments.

The PCL is a widely used questionnaire to assess the severity of traumatic reaction and to screen those with a PTSD diagnosis. There are several versions of the PCL, including the PCL-Civilian version, PCL-Military version for people with combat experience, and PCL-Specific version for people who have experienced specific traumatic events. Its psychometric and screening properties have been well reported (McDonald & Calhoun, 2010; Wilkins, Lang, & Norman, 2011), and there are several abbreviated versions to improve clinical utility (Bliese et al., 2008; Lang & Stein, 2005). The PCL has also been used among traumatized Japanese populations (Sakuma et al., 2015; Yabe et al., 2014; Yasumura et al., 2012); however, the Japanese version of the PCL has yet to be validated.

Accordingly, the aims of this study were to (1) examine the psychometric property of the Japanese version of the PCL-S and (2) compare the diagnostic accuracy of the PCL-S with that of the IES-R as well as the abbreviated version of the PCL-S among the evacuees of the Fukushima Daiichi Nuclear Power Plant (NPP) accident.

1 | METHODS

1.1 | Participants

To recruit people with a range of traumatic reaction levels, we included evacuee and clinical participants. The inclusion criteria for evacuee participants were people who (1) used to live within the government-designated evacuation zone, (2) responded to the Mental Health and Lifestyle Survey of the Fukushima Health Management Survey conducted in 2013 (Yasumura et al., 2012), and (3) were at least 16 years old. The candidates were selected on the basis of Kessler's 6 items for psychological distress (K6) (Kessler et al., 2003; Sakurai, Nishi, Kondo, Yanagida, & Kawakami, 2011) and PCL-S scores in the survey, 10 each from the low-, middle-, and high-score categories. The inclusion criteria for clinical participants were patients who (1) visited the Department of Psychiatry of Fukushima Medical University Hospital and its related institutions, (2) received a clinical diagnosis of PTSD or adjustment disorder from the attending psychiatrist, (3) were permitted to participate in this study by the psychiatrist, and (4) were at least 16 years old. In total, 38 evacuee participants and 13 clinical participants were recruited.

1.2 | Procedure

Participants were asked to fill in the self-administered PCL-S and IES-R, followed by a structured interview using the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). The second PCL-S was administered after 1 week by mail to examine test-retest reliability.

1.3 | Screening instruments

The PCL is a self-administered questionnaire assessing the 17 symptoms of PTSD on the basis of the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) (American Psychiatric Association, 1994), which includes 3 symptom clusters: re-experiencing, avoidance/numbing, and arousal. Participants indicated whether they were bothered by symptoms due to the traumatic event in the past month on a 5-point Likert scale (1 = not at all to 5 = extremely), with the sum of the score ranging from 17 to 85. We used the PCL-S, with the Great East Japan Earthquake—including the earthquake, tsunami, and NPP accident—specified as the traumatic event.

The original PCL has a Cronbach alpha of 0.939, and its correlation with the Clinician-Administered PTSD Scale for DSM-IV is 0.929. The sensitivity and specificity for PTSD diagnosis are 0.778 and 0.864, respectively, with a cutoff point of 49/50, and 0.944 and 0.864, respectively, with a cutoff point of 43/44 among motor vehicle accident survivors or survivors of sexual assault in the United States (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996).

After the English-Japanese translation was authorized by the original author of the scale, a Japanese psychiatrist translated the original English version of the PCL-S into Japanese, and then it was back-translated by 2 native English-speaking bilingual scientists. The back-translated version was then compared with the original, and adjustments were made to the Japanese version, considering linguistic and semantic equivalents.

There are 2 evaluation methods: the cutoff point method for the total sum of the 17 items and the symptom cluster method (SCM), which requires 1 re-experiencing, 3 numbing/avoidance, and 2 hyperarousal symptoms according to the DSM-IV. The symptom is regarded as present for scores of 3 or more, representing at least moderately bothersome symptoms, as well as for scores of 4 or more, representing at least quite a bit symptoms, assuming that Japanese people tend to present psychological symptom less (Harada et al., 2012).

Abbreviated versions of the PCL have been proposed, and we chose to examine the following 3 versions for brevity and for optimal diagnostic utility: Bliese's 4 items, which include (1) intrusive recollections, (5) reaction to reminders, (7) avoid reminders, and (15) concentration difficulties (Bliese et al., 2008), as well as Lang and Stein's 4 and 6 items (Lang & Stein, 2005). Lang and Stein's 4 items include (1) intrusive recollections, (4) distress at reminders, (7) avoid reminders, and (16) hypervigilance. The 6 items include (1) intrusive recollections, (4) distress at reminders, (7) avoid reminders, (10) detached from others, (14) irritability/anger, and (15) concentration difficulties.

The IES-R is a self-administered questionnaire on 22 traumatic symptoms rated on a 5-point Likert scale (0-4) (Weiss & Marmar, 1997). The total scores range from 0 to 88 with higher scores representing greater severity. The Japanese version of the IES-R has been validated (Asukai et al., 2002). Cronbach alphas for the subscales are 0.86 to 0.91 for intrusion, 0.81 to 0.90 for avoidance, and 0.80 to 0.86 for hyperarousal (Weiss, 2004). Although the IES-R was not developed for making categorical PTSD diagnosis, various cutoff points have been proposed to indicate probable PTSD, with a range from 19 to 35 (Asukai et al., 2002; Bienvenu, Williams, Yang, Hopkins, & Needham, 2013; Chen, Cheng, & Yen, 2011; Creamer, Bell, & Failla, 2003).

1.4 | Reference standard

Posttraumatic stress disorder diagnosis was made using the PTSD section of the WHO-CIDI (Kessler & Ustun, 2004). This structured interview was conducted by 6 health professionals who underwent interview training. The interviewers were blind to the clinical diagnosis.

1.5 | Analysis

We analyzed the data of participants who responded to the PCL-S without missing answers (n = 48). For test-retest reliability, we examined only those who completed the PCL-S at both time points (n = 33). All participants experienced at least 1 event of the disaster, and thus, we included all in the analysis.

First, to examine psychometric properties, we calculated Cronbach alpha to evaluate the reliability of the PCL-S. We then calculated the Spearman rank-order correlation of the PCL-S scores 1 week apart to examine test-retest reliability. We also calculated the Spearman rank-order correlation between the PCL-S and IES-R to examine concurrent validity. Then, to examine diagnostic accuracy of the PCL-S and IES-R for PTSD diagnosis over the past 30 days on the basis of the WHO-CIDI, we calculated sensitivity, specificity, and diagnostic efficiency, which is the proportion of those correctly categorized as true-positive and true-negative. Posttraumatic stress disorder diagnosis was made according to the DSM-IV and International Classification of Diseases, 10th revision (ICD-10) (WHO, 1993). Area under receiver operating characteristic (ROC) curves (AUCs) and their 95% confidence intervals (CIs) were calculated, and the optimal cutoff point was examined using the Youden method (Fluss, Faraggi, & Reiser, 2005). Similarly, the screening properties for the abbreviated versions of the PCL-S were examined. All statistical analyses were performed using Stata 13.0 for Windows (StataCorp LP, College Station, Texas).

1.6 | Ethical consideration

The study was approved by the Ethics Committees of the Fukushima Medical University (numbers 1316 and 1489) and National Center of Neurology and Psychiatry (A2014-160). After informing participants that their participation was voluntary, that they could withdraw from the study at any time, and that they would not be disadvantaged in any way if they chose to withdraw or decline to participate, receipt of a returned questionnaire was assumed to indicate consent for the Mental Health and Lifestyle Survey of the Fukushima Health Management Survey, and written consent was obtained for the diagnostic study. Authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

2 | RESULTS

2.1 | Participants' characteristics

This study was conducted from November 2013 to March 2014. The participants who met the diagnostic criteria for PTSD in the past 30 days were 15 (31.3%) by the DSM-IV and 14 (29.2%) by the ICD-10 (Table 1). Among the participants, 24 (50.0%) were above the conventional cutoff points of 44, and 19 (39.6%) were above the cutoff of 50. A comparison of PCL-S scores by the experience of the Great East Japan Earthquake, other traumatic events, and functional impairment is presented in Table 2. There were no associations between PCL-S scores and the experience of earthquake, tsunami, NPP accident, or life-threatening experience during the Great East Japan Earthquake or another traumatic event. The PCL-S scores were higher among those who reported functional impairment than among those who did not (median score: 50 vs 35, respectively, $z = 3.1$, $P = .002$).

2.2 | Psychometric properties of the PCL-S

Cronbach alpha of the PCL-S was 0.92 for all 17 items and 0.83 for re-experiencing, 0.82 for avoidance/numbing, and 0.79 for hyperarousal. For test-retest reliability, the mean score (SD) was 42.4 (15.0) for the first test and 41.2 (15.7) for the second, with a difference of 1.27 ($t = 0.860$, $P = .396$). The Spearman rank-order correlation was 0.85 ($P < .001$). The Spearman rank-order correlation between PCL and Japanese version IES-R scores was 0.90 ($P < .001$) among those who completed both scales (n = 47).

2.3 | Diagnostic accuracy of the PCL-S

The flow of participants who underwent the PCL-S and subsequent PTSD diagnosis (past 30 days) according to the DSM-IV is presented in Figure 1. The median PCL-S score was higher among those with PTSD than among those without (58 and 36, respectively).

The indicators of the diagnostic accuracy of the PCL-S and IES-R are presented in Table 3. The AUC was 0.83 (95% CI, 0.71-0.95) for the DSM-VI and 0.79 (95% CI, 0.65-0.92) for the ICD-10, suggesting moderate accuracy for both. The optimal cutoff point was 52 for the DSM-VI and 46 for the ICD-10. In reference to the IES-R, AUC was 0.70 (95% CI, 0.56-0.85) for the DSM-VI and 0.75 (95% CI, 0.60-0.89) for the ICD-10. The ROC curves of the PCL and IES-R for PTSD diagnosis based on the DSM-IV and ICD-10 are presented in Figure 2.

Regarding the SCM for the PCL-S, the AUC was 0.68 (95% CI, 0.53-0.83) for the DSM-VI with assumption of 3 or above as symptom present and 0.70 (95% CI, 0.57-0.84) with assumption of 4 or above as symptom present. Agreement on PTSD diagnosis was 68.8% ($\kappa = 0.33$, SE = 0.14, $z = 2.37$, $P = .009$) between SCM (3+) and the DSM-IV and 79.2% ($\kappa = 0.46$, SE = 0.14, $z = 3.34$, $P < .001$) between the SCM (4+) and the DSM-IV.

The details on the screening properties for DSM-IV-based PTSD diagnosis of the 3 abbreviated versions of the PCL-S at its optimal cutoff points are presented in Table 3. The AUCs were 0.86 (95% CI, 0.75-0.98) for Bliese's 4 items at the cutoff point of 12 and 0.82 (95% CI: 0.70–0.95) for Lang and Stein's 4 items at the cutoff point of 13. The AUC for the 6 items proposed by Lang and Stein was 0.85 (95% CI, 0.73-0.97) for the DSM-VI at the cutoff point of 17. The ROC curves for the abbreviated version of the PCL-S are presented in Figure 3.

3 | DISCUSSION

The psychometric properties of the Japanese version of the PCL-S showed satisfactory internal consistency and very strong correlation in examining test-retest reliability, and concurrent validity with IES-R. The Japanese version of the PCL-S demonstrated moderate diagnostic accuracy and improved performance over the IES-R for DSM-IV-based PTSD diagnosis for the past 30 days. The cutoff point method for PCL performed better than did the SCM.

TABLE 1 Gender, age, and proportion of PTSD of the participants

		Overall		Evacuee (n = 35)		Clinical (n = 13)	
		n or Mean	% or SD	n or Mean	% or SD	n or Mean	% or SD
Gender	Men	23	47.9	16	45.7	7	53.9
	Women	25	52.1	19	54.3	6	46.2
Age, mean, SD		62.5	14.8	66.6	11.6	51.5	17.2
PTSD diagnosis							
Past 30 days	DSM-IV	15	31.3	9	25.7	6	46.2
	ICD-10	14	29.2	9	25.7	5	38.5
Past 12 months	DSM-IV	17	35.4	11	31.4	6	46.2
	ICD-10	18	37.5	13	37.1	5	38.5
Lifetime	DSM-IV	20	41.7	14	40.0	6	46.2
	ICD-10	21	43.8	16	45.7	5	38.5
PCL	44+	24	50.0	19	54.3	5	38.5
	50+	19	39.6	14	40.0	5	38.5

Abbreviations: DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition; ICD-10, International Classification of Diseases, 10th revision, PCL, PTSD Checklist; PTSD, posttraumatic stress disorder.

TABLE 2 Comparison of PCL-S scores by experience of the Great East Japan Earthquake, other traumatic events, and functional impairment

	n	Median	25th, 75th Percentiles	z^a	*P*
Overall	48	43.5	34, 53		
Experience of the Great East Japan Earthquake					
Earthquake					
Yes	42	44.5	34, 55	1.3	0.201
No	6	38.5	26, 43		
Tsunami					
Yes	21	43	34, 58	0.2	0.827
No	27	44	36, 53		
NPP accident					
Yes	46	43.5	34, 53	−0.4	0.661
No	2	46	39, 53		
Life-threatening experience during the Great East Japan Earthquake[b]					
Yes	34	44.5	36, 58	1.3	0.212
No	13	36	30, 50		
Traumatic experience other than the Great East Japan Earthquake[b]					
Yes	9	50	44, 53	1.3	0.204
No	38	40	30, 53		
Functional impairment[c]					
Yes	29	50	38, 58	3.1	0.002
No	19	35	22, 45		

Abbreviations: NPP, nuclear power plant; PCL-S, PTSD Checklist-Specific.

[a]Mann-Whitney *U* test.

[b]n = 47 due to 1 missing observation.

[c]Yes: often, sometimes; No: rarely, never.

3.1 | Psychometric properties of PCL-S

The internal consistency of the Japanese version of the PCL-S was satisfactory, and the Cronbach alpha of 0.92 was comparable with a previous report in terms of both the overall scale and its subscales (Wilkins et al., 2011). The test-retest reliability at 1 week was 0.90, which fell in the range of 0.68 to 0.92 observed in previous studies (Wilkins et al., 2011). Concurrent validity was confirmed, as demonstrated by the Spearman rank-order correlation of 0.90 between the total scores of the PCL-S and IES-R. Overall, the Japanese version of the PCL was demonstrated to be reasonably reliable and valid.

3.2 | Diagnostic accuracy of PCL-S

On the basis of the ROC curves, we determined that the optimal cutoff points of the Japanese version of the PCL-S for the past 30 days' PTSD

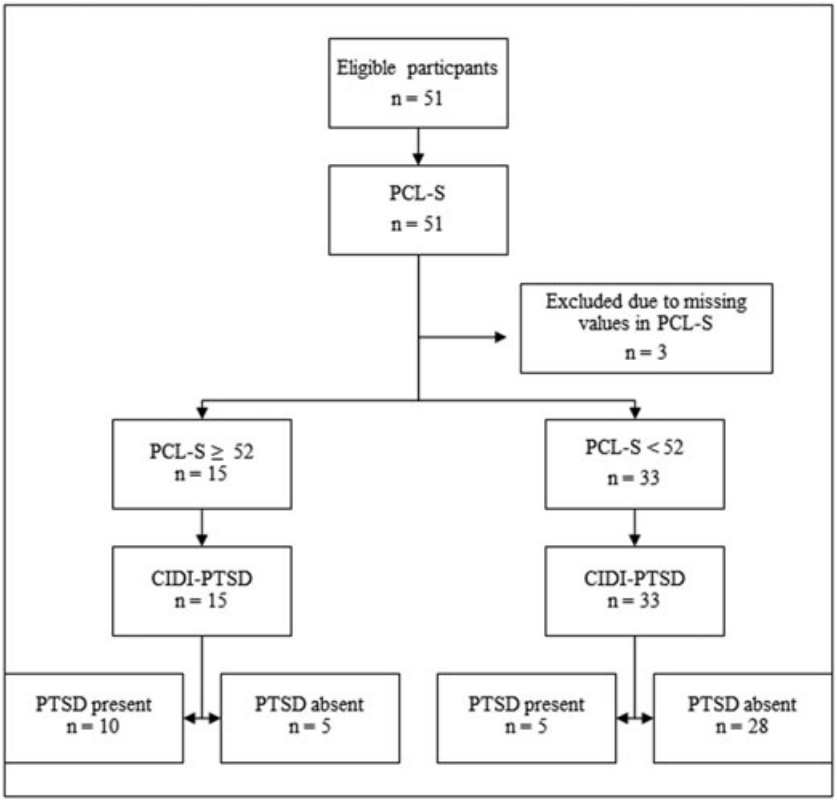

FIGURE 1 Flowchart of the participants and assessment results by PCL-S and past 30 days' PTSD diagnosis made by WHO-CIDI. PCL-S: PTSD Checklist-Specific. PTSD: posttraumatic stress disorder. WHO-CIDI: World Health Organization Composite International Diagnostic Interview

TABLE 3 Screening properties for PTSD diagnosis of the PCL-S and IES-R among evacuees of the Fukushima NPP accident

	n	ROC Area	95%CI Lower Limit	95%CI Upper Limit	Optimal Cutoff	Sensitivity (%)	Specificity (%)	Diagnostic Efficiency (%)	LR+	LR−
Cut point method for PCL-S and IES-R										
PCL-S total score										
DSM-IV	48	0.83	0.71	0.95	52	66.7	84.9	79.2	4.40	0.39
ICD-10	48	0.79	0.65	0.92	46	78.6	70.6	72.9	2.67	0.30
IES-R										
DSM-IV	49	0.70	0.56	0.85	37	73.3	62.5	66.0	1.96	0.43
ICD-10	49	0.75	0.60	0.89	37	78.6	63.6	68.1	2.16	0.34
Symptom cluster method for PCL-S										
3+ on the Likert scale as symptom present										
DSM-IV	48	0.68	0.53	0.83	1	66.7	69.7	68.8	2.20	0.48
4+ on the Likert scale as symptom present										
DSM-IV	48	0.70	0.57	0.84	1	46.7	93.9	79.2	7.70	0.57
Abbreviated versions										
Bliese's 4 items										
DSM-IV	48	0.86	0.75	0.98	12	73.3	84.9	81.3	4.84	0.31
Lang and Stein's 4 items										
DSM-IV	48	0.82	0.70	0.95	13	60.0	87.9	79.2	4.95	0.46
Lang and Stein's 6 items										
DSM-IV	48	0.85	0.73	0.97	17	80.0	75.8	77.1	3.30	0.26

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition; IES-R, Impact of Event Scale-Revised; ICD-10, International Classification of Diseases, 10th revision; LR−, negative likelihood ratio; LR+, positive likelihood ratio; PCL-S, PCL-Specific; PTSD, posttraumatic stress disorder.

An optimal cut point of 1 means PTSD diagnosis is present according to the symptom cluster method.

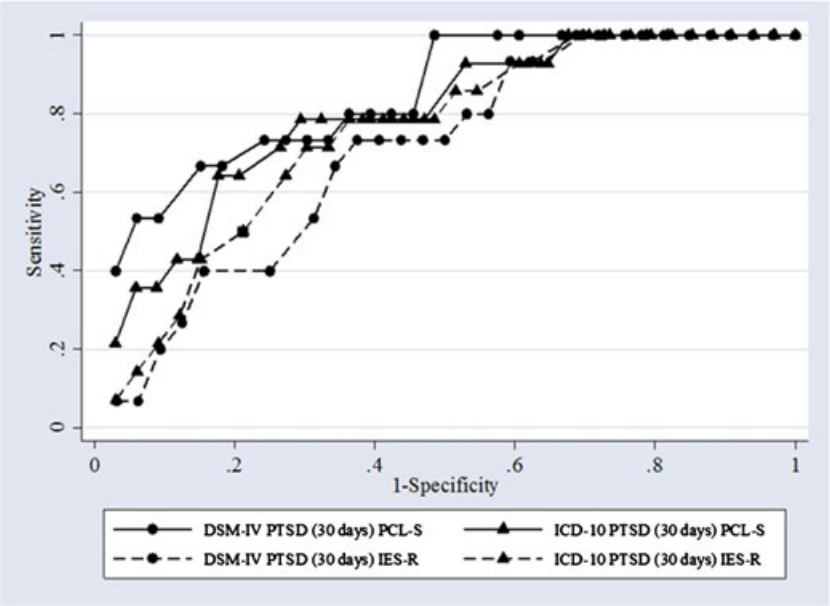

FIGURE 2 The ROC curves of PCL-S and IES-R scores for the past 30 days' PTSD diagnosis based on the DSM-IV and ICD-10. ROC: receiver operating characteristic; PCL-S: PTSD Checklist-Specific; IES-R: Impact of Event Scale-Revised; PTSD: posttraumatic stress disorder; DSM-IV: *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition; ICD-10: International Classification of Diseases, 10th revision

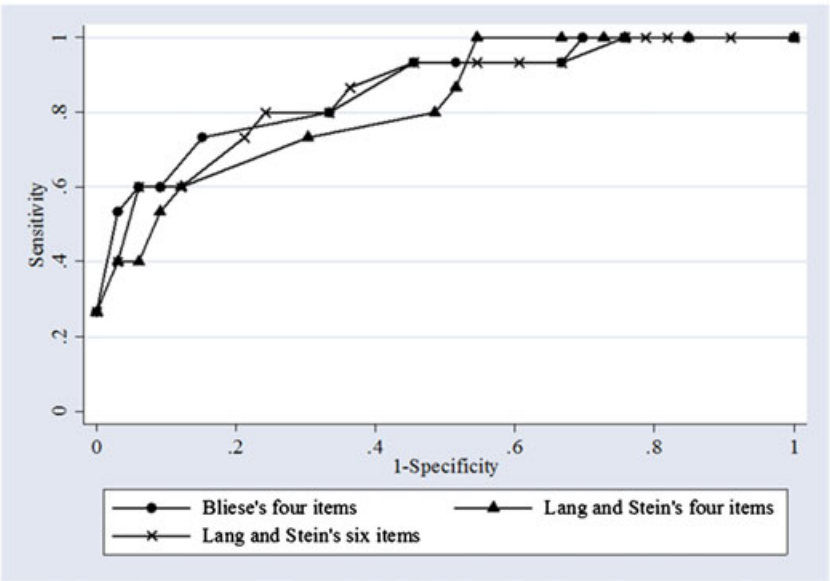

FIGURE 3 The ROC curves of the 3 abbreviated versions of the PCL-S for the past 30 days' PTSD diagnosis based on the DSM-IV. ROC: receiver operating characteristic; PCL-S: PTSD Checklist-Specific; PTSD: posttraumatic stress disorder; DSM-IV: *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition

diagnosis were 52 for the DSM-IV and 46 for the ICD-10 among individuals who experienced the Great East Japan Earthquake and Fukushima NPP accident. Our result lies on the higher end of the reported score range from 32 to 50, which varies depending on the study population and type of trauma exposure (McDonald & Calhoun, 2010). The screening properties—sensitivity of 66.7% and specificity of 84.9% at a cutoff point of 52—fall within the range observed in previous studies, which have found sensitivity to be 60% to 94% and specificity to be 86% to 99% for the PCL-S (McDonald & Calhoun, 2010).

Traditionally, the optimal cutoff point has been determined on the basis of the ROC curve by balancing sensitivity and specificity; however, this approach has recently come into question (Wald & Bestwick, 2014). The optimal cut point should be examined depending on the intended purpose of the use (McDonald & Calhoun, 2010). For example, diagnostic efficiency was relatively high (79.2%) at the cutoff point of 52 based on the ROC curve, with a low sensitivity (66.7%) and high specificity (84.9%). To more broadly capture people at risk of PTSD following a complex disaster in a community, a lower cutoff point is desirable, as it increases the sensitivity. Then, further detailed assessment is needed to confirm the diagnosis.

In comparing the diagnostic accuracy of PCL-S and IES-R, we found that the PCL-S was superior to the IES-R in detecting the past 30 days' PTSD according to the DSM-IV definition. On the other hand, at the cutoff point of 37, the IES-R performed better in discriminating PTSD cases and non-cases according to the ICD-10 definition than it did by the DSM-IV definition. The IES-R was originally developed to measure degree of traumatic distress, not to diagnose PTSD. Nevertheless, the IES-R may have performed better in screening cases of PTSD according to the ICD-10, as PTSD is operationalized as a broader concept compared with the definition in the DSM-IV (Van Ameringen, Mancini, & Patterson, 2011). The cutoff point of 37 was higher than the previously proposed cutoff points of 19 to 35. Our study may have demonstrated a higher cutoff point because of the different trauma and time since exposure.

The cutoff point method performed better than did the SCM, as the AUC was 0.83 for the cutoff point of 52 and 0.68 for the SCM. Interestingly, if we assume the presence of symptoms for responses of 4 points or more, performance was better than it was when assuming symptoms at 3 points or more. Kappa was higher with symptom presence at 4 points or more. These improvements resulted from increased specificity when 4 points indicated symptom presence. As the participants presented with a high degree of traumatic distress, a higher threshold for determining symptom presence may have decreased the number of false-positives, resulting in higher specificity. To increase specificity in detailed secondary assessment, the use of the SCM with 4 points or more indicating symptom presence may be preferable.

The results supported the use of an abbreviated version of the Japanese version of PCL-S, for both the 4- and 6-item versions, as the screening properties were comparable or even better than those of the full PCL-S. The best cutoff point was 12 for Bliese's 4 items and 17 for Lang and Stein's 6 items, and each was higher than the previous report of 7 and 14, respectively. In the aftermath of a disaster, the use of Lang and Stein's 4 items was tested among the people affected by Hurricane Katrina, but this usage was not validated (Hirschel & Schulenberg, 2010). In our study, the sensitivity of Lang and Stein's 4 items was 60.0%, which was the lowest of the abbreviated versions, and this may not be appropriate to broadly capture those at high risk of PTSD diagnosis. Further studies on abbreviated versions of the PCL-S are needed to draw conclusions, as there are limited empirical studies on the abbreviated versions.

3.3 | Limitations

There are several limitations of this study. Although we recruited participants with different degrees of traumatic reactions, the sample size was relatively small. Specifically, we could not recruit the targeted number of clinical participants, because there were few patients with the diagnosis of PTSD at medical institutions for unknown reason. A further validation study with a larger and more representative sample is warranted. Second, although we recruited people who experienced the Great East Japan Earthquake, the symptoms measured in this study reflect not only the traumatic event but also, and perhaps more largely, secondary stressors after the disaster, as suggested by previous research (Lock et al., 2012). This concern is supported by the finding that there was no difference in PCL-S scores by experience of disaster or life-threatening experience. The relationship between reaction to traumatic events and secondary life stressors should be differentiated in further studies. Lastly, the diagnostic criteria of PTSD have changed with the introduction of the DSM-5 (American Psychiatric Association, 2013), and the revision of ICD-10 will follow shortly. The use of PCL should be examined with this dynamic context in mind.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Asukai, N., Kato, H., Kawamura, N., Kim, Y., Yamamoto, K., Kishimoto, J., … Nishizono-Maher, A. (2002). Reliability and validity of the Japanese-language version of the Impact of Event Scale-Revised (IES-R-J), four studies of different traumatic events. *The Journal of Nervous and Mental Disease*, *190*(3), 175–182.

Bienvenu, O. J., Williams, J. B., Yang, A., Hopkins, R. O., & Needham, D. M. (2013). Posttraumatic stress disorder in survivors of acute lung injury: Evaluating the Impact of Event Scale-Revised. *Chest*, *144*(1), 24–31.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, *34*(8), 669–673.

Bliese, P. D., Wright, K. M., Adler, A. B., Cabrera, O., Castro, C. A., & Hoge, C. W. (2008). Validating the primary care posttraumatic stress disorder screen and the posttraumatic stress disorder checklist with soldiers returning from combat. *Journal of Consulting and Clinical Psychology*, *76*(2), 272–281.

Chen, C. S., Cheng, C. P., & Yen, C. F. (2011). Validation of the Impact of Event Scale-Revised for adolescents experiencing the floods and mudslides. *Kaohsiung Journal of Medical Sciences*, *27*(12), 560–565.

Creamer, M., Bell, R., & Failla, S. (2003). Psychometric properties of the Impact of Event Scale-Revised. *Behaviour Research and Therapy*, *41*(12), 1489–1496.

Fluss, R., Faraggi, D., & Reiser, B. (2005). Estimation of the Youden index and its associated cutoff point. *Biometrical*, *47*(4), 458–472.

Harada, N., Takeshita, J., Ahmed, I., Chen, R., Petrovitch, H., Ross, G. W., & Masaki, K. (2012). Does cultural assimilation influence prevalence and presentation of depressive symptoms in older Japanese American men? The Honolulu-Asia aging study. *The American Journal of Geriatric Psychiatry*, *20*(4), 337–345.

Hirschel, M. J., & Schulenberg, S. E. (2010). On the viability of PTSD Checklist (PCL) short form use: Analyses from Mississippi Gulf Coast Hurricane Katrina survivors. *Psychological Assessment*, *22*(2), 460–464.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., … Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*(2), 184–189.

Kessler, R. C., & Ustun, T. B. (2004). The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 93–121.

Lang, A. J., & Stein, M. B. (2005). An abbreviated PTSD checklist for use as a screening instrument in primary care. *Behaviour Research and Therapy*, *43*(5), 585–594.

Lock, S., Rubin, G. J., Murray, V., Rogers, M. B., Amlot, R., & Williams, R. (2012). Secondary stressors and extreme events and disasters: A systematic review of primary research from 2010 to 2011. *PLoS Current*, *4*. 10.1371/currents.dis.a9b76fed1b2dd5c5bfcfc13c87a2f24f.

McDonald, S. D., & Calhoun, P. S. (2010). The diagnostic accuracy of the PTSD checklist: A critical review. *Clinical Psychology Review*, *30*(8), 976–987.

Neria, Y., Nandi, A., & Galea, S. (2008). Post-traumatic stress disorder following disasters: A systematic review. *Psychological Medicine*, *38*(4), 467–480.

North, C. S., & Pfefferbaum, B. (2013). Mental health response to community disasters: A systematic review. *JAMA*, *310*(5), 507–518.

Sakuma, A., Takahashi, Y., Ueda, I., Sato, H., Katsura, M., Abe, M., … Matsumoto, K. (2015). Post-traumatic stress disorder and depression prevalence and associated risk factors among local disaster relief and reconstruction workers fourteen months after the Great East Japan Earthquake: A cross-sectional study. *BMC Psychiatry*, *15*, 58. 10.1186/s12888-015-0440-y.

Sakurai, K., Nishi, A., Kondo, K., Yanagida, K., & Kawakami, N. (2011). Screening performance of K6/K10 and other screening instruments for mood and anxiety disorders in Japan. *Psychiatry and Clinical Neurosciences*, *65*(5), 434–441.

Terhakopian, A., Sinaii, N., Engel, C. C., Schnurr, P. P., & Hoge, C. W. (2008). Estimating population prevalence of posttraumatic stress disorder: An example using the PTSD checklist. *Journal of Traumatic Stress*, *21*(3), 290–300.

Van Ameringen, M., Mancini, C., & Patterson, B. (2011). The impact of changing diagnostic criteria in posttraumatic stress disorder in a Canadian epidemiologic sample. *The Journal of Clinical Psychiatry*, *72*(8), 1034–1041.

Wald, N. J., & Bestwick, J. P. (2014). Is the area under an ROC curve a valid measure of the performance of a screening or diagnostic test? *Journal of Medical Screening*, *21*(1), 51–56.

Weathers, F. W., Litz, B. T., Herman, D. S., Hushka, J. A. & Keane, T. M. (1993). The PTSD Checklist (PCL), reliability, validity, and diagnostic utility. *The Annual Meeting of International Society for Traumatic Stress Studies*. San Antonio, TX.

Weiss, D. S. (2004). The Impact of Event Scale-Revised. In J. P. Wilson, & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (). (pp. 168–189). New York: Guilford Press.

Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale-Revised. In J. P. Wilson, & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (). (pp. 399–411). New York: Guilford Press.

Wilkins, K. C., Lang, A. J., & Norman, S. B. (2011). Synthesis of the psychometric properties of the PTSD checklist (PCL) military, civilian, and specific versions. *Depression and Anxiety*, *28*(7), 596–606.

World Health Organization (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. Geneva: World Health Organization.

Yabe, H., Suzuki, Y., Mashiko, H., Nakayama, Y., Hisata, M., Niwa, S., … Abe, M. (2014). Psychological distress after the Great East Japan Earthquake and Fukushima Daiichi Nuclear Power Plant accident: Results of a mental health and lifestyle survey through the Fukushima Health Management Survey in FY2011 and FY2012. *Fukushima Journal of Medical Science*, *60*(1), 57–67.

Yasumura, S., Hosoya, M., Yamashita, S., Kamiya, K., Abe, M., Akashi, M., … Ozasa, K. (2012). Study protocol for the Fukushima Health Management Survey. *Journal of Epidemiology*, *22*(5), 375–383.

How to cite this article: Suzuki, Y., Yabe, H., Horikoshi, N., Yasumura, S., Kawakami, N., Ohtsuru, A., Mashiko, H., Maeda, M., and on behalf of the Mental Health Group of the Fukushima Health Management Survey (2016), Diagnostic accuracy of Japanese posttraumatic stress measures after a complex disaster: the Fukushima Health Management Survey, *Asia-Pacific Psychiatry*, doi: 10.1111/appy.12248