Journal of Affective Disorders 121 (2010) 212–219

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research report

Validation of the Center for Epidemiologic Studies Depression Scale in screening for major depressive disorder among retired firefighters exposed to the World Trade Center disaster

Sydney Chiu [a], Mayris P. Webber [a,b,*], Rachel Zeig-Owens [a], Jackson Gustave [a], Roy Lee [a], Kerry J. Kelly [a], Linda Rizzotto [a], David J. Prezant [a,b]

[a] Fire Department of the City of New York, Bureau of Health Services, Brooklyn, New York 11201, United States
[b] Montefiore Medical Center and Albert Einstein College of Medicine, Bronx, New York 10461, United States

ARTICLE INFO

Article history:
Received 22 April 2009
Received in revised form 28 April 2009
Accepted 21 May 2009
Available online 18 June 2009

Keywords:
Depression
Center for Epidemiologic Studies Depression Scale
Validation
World Trade Center
Firefighter

ABSTRACT

Background: We evaluated the performance of a modified Center of Epidemiologic Studies Depression Scale (CES-D-m), which captured symptoms in the past month, in comparison to the Diagnostic Interview Schedule (DIS) in identification of major depressive disorder (MDD) in World Trade Center (WTC)-exposed retired Fire Department, City of New York (FDNY) firefighters. *Methods:* From 12/2005 to 7/2007, FDNY enrolled retired firefighters in its Medical Monitoring and Treatment Program. All participants completed the CES-D-m and the DIS on the same day. Sensitivity, specificity, receiver operating characteristic (ROC) curves, and Youden's index were used to assess properties of the CES-D-m. Multivariate logistic regression analyses were also used. *Results:* 7% of 1915 retired male firefighters were diagnosed with MDD using the DIS. Using the most common CES-D cutoff score of 16, the prevalence of elevated risk was 36%, which declined to 23% using a cutoff score of 22, as determined by Youden's index. At 22, CES-D-m sensitivity was 0.84, specificity was 0.82, and the area under the ROC curve was 0.89 relative to DIS MDD diagnosis. *Limitations:* Participants were more likely than non-participants to live in the New York City area. *Conclusions:* This is the first study of WTC rescue/recovery workers to assess the performance of a one-month version of the CES-D. The CES-D-m performed well in identifying those at elevated risk. Since diagnostic follow-up is time consuming and costly, it is important to correctly distinguish those at elevated risk using a screening tool that has been validated in the population under study.

1. Introduction

On September 11, 2001 (9/11), approximately 11,500 Fire Department, City of New York (FDNY) firefighters responded to the terrorist attacks on the World Trade Center (WTC). The rescue and recovery effort exposed many of the responders to physical and emotional trauma, putting them at risk for various mental health disorders including major depressive disorder (MDD). Although the etiology of MDD is not clearly understood (Person et al., 2006), several studies performed on specific post-disaster populations (North et al., 1999; Tak et al., 2007), including 9/11 (Galea et al., 2002; Gross et al., 2006; Long et al., 2007; Stellman et al., 2008; Tapp et al., 2005), highlight the concern of experts in the mental health community over depressive symptoms following traumatic events (Katz et al., 2002). A study conducted in October and November 2001 of Manhattan residents living less than 1 mile from the WTC reported 16.8% with depression-like symptoms (Galea et al., 2002). In addition, 12% of a sample of NYC transit workers stationed near the WTC towers screened in 2002 (Tapp et al., 2005) and 16.1% of a sample of WTC clean up and recover workers screened in 2003 reported symptoms consistent with major depression (Gross et al., 2006).

Given the prevalence of and health risks associated with depression, self-administered screening tools provide practical

* Corresponding author. New York City Fire Department, Bureau of Health Services, 9 Metrotech Center, Brooklyn, New York 11201, United States. Tel.: +1 718 999 2665; fax: +1 718 999 0142.
E-mail address: webberm@fdny.nyc.gov (M.P. Webber).

0165-0327/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jad.2009.05.028

means for assessing depressive symptoms. Screening tools identify persons who reach a predetermined threshold or cutoff score for follow-up with more intensive and time-consuming diagnostic tests. The Center for Epidemiologic Studies Depression Scale (CES-D) was designed in 1977 to screen for depressive symptoms in the general population. It consists of 20 self-reported items, rated on a 4-point scale of symptom frequency, and is scored from 0 to 60 (Radloff, 1977), with higher scores more likely to be associated with a MDD diagnosis. While the original work reported an optimum cutoff score of 16 in the general population, others have reported cutoff scores from 12 (Watson et al., 2004) to 25 (Haringsma et al., 2004) in specific populations, with 16 most frequently used (Mair et al., 2009; Schiepers et al., 2009; Shinar et al., 1986; Snoek et al., 2008).

Following administration of a screening instrument, those who exceed a specified threshold are advised to complete a follow-up diagnostic interview for a more comprehensive review of MDD symptoms. The Diagnostic Interview Schedule (DIS) is a structured diagnostic interview designed to assist in making a MDD diagnosis according to Diagnostic and Statistical Manual for Mental Disorders, version IV (DSM-IV-TR®) criteria (Robins et al., 1995). Full assessment of accepted diagnostic criteria (such as with the DIS or clinical interviews) is required for diagnosis of MDD.

The DSM-IV-TR® identifies nine symptoms that are associated with a diagnosis of MDD and requires that at least five of these nine symptoms be present over a two-week period. In contrast, the CES-D assesses symptoms only during the past week (Radloff, 1977) and therefore encompasses a time frame that is not consistent with the DSM-IV-TR® definition of MDD.

To our knowledge, no study has determined the optimal cutoff score for depression using the CES-D or any other tool in a population of rescue/recovery workers after a major disaster. The purpose of the current study was to evaluate the performance characteristics of the CES-D modified to encompass a one-month assessment of depressive symptoms (CES-D-m) in relation to a full diagnostic assessment using a structured diagnostic interview (the DIS) in a large retired population of WTC-exposed firefighters. Further, since the FDNY is currently using a cutoff score of 16, which means that only persons scoring at or above that level receive a full diagnostic assessment, our intent was to evaluate the screening efficiency of this cutoff.

2. Methods

Since 1997, the FDNY Bureau of Health Services (BHS) performs periodic health evaluations on active FDNY members approximately every 18 months, which include both physician examinations and, since 2001, self-administered health questionnaires. In 2005, a more extensive mental health questionnaire incorporated the CES-D. This occurred co-incident with outreach to retired WTC-exposed firefighters, who were invited to resume participation in health screenings, which were previously limited to the active workforce. During the first 19 months of the retiree screenings (12/2005–7/2007) all retired firefighters completed both the CES-D-m and the DIS on the same day, regardless of whether a threshold score was attained. Participation in the study required written informed consent and was approved by the institutional review board of Montefiore Medical Center.

2.1. Study participants

Eligibility for study enrollment included retired FDNY employees who worked at least one shift at any of the designated WTC work sites between 9/11/2001 and 7/25/2002, when the site was formally closed. Active recruitment involved contacting retirees through invitational letters and/or by telephone using contact information from the FDNY pension database. Passive recruitment involved self-enrollment through the FDNY website. There were 4080 eligible candidates for WTC MMTP retiree enrollment.

By the end of the initial enrollment period, 2574 retired firefighters (63%) had returned to FDNY BHS and completed a monitoring exam including the mental health questionnaire. We excluded the following: persons who could not be classified in a 9/11 exposure group or first arrived at any of the WTC sites after 9/24/2001 ($n=316$); those who did not complete the CES-D-m and the DIS on the same day ($n=218$); fire marshals ($n=84$), due to the distinct nature of their work; persons who retired due to a mental health disability ($n=31$); and females ($n=10$), because they represent a small proportion of the workforce. The final analytic population totaled 1915 (47% of eligible).

The active firefighter population was analyzed for comparison purposes. After exclusions identical to those in retirees, the active population totaled 4543 who took the CES-D-m from 12/2006 to 12/2007. Since active firefighters completed the DIS only when their CES-D-m scores equaled or exceeded a cutoff score of 16, only CES-D-m responses from active firefighters were analyzed for comparison with those of retirees.

2.2. Data sources

Age, hire date, retirement date and status (ordinary vs. disability for any cause other than psychological), and rank (chiefs, captains, lieutenants, and firefighters) were obtained from the FDNY employee database. All other variables were collected from the questionnaires.

2.3. 9/11 exposure

The FDNY-WTC Exposure Intensity Index (Prezant et al., 2002) categorized exposure based on first arrival at the WTC site as follows: “Group 1”, the most severely exposed, arriving on the morning of 9/11 and present during the tower collapses; “Group 2” arriving during the afternoon of 9/11; “Group 3” arriving the next day on 9/12/2001; and “Group 4”, or least exposed, arriving any day between days 3 and 14.

2.4. Measures

The CES-D-m was incorporated into the mental health questionnaire to identify persons requiring further evaluation. The 20 item questionnaire assesses depressive symptoms; each symptom is scored on a scale of 0 to 3 (0 = “None or rarely [less than 1 day per week]” to 3 = “Most or all the time

[5–7 days per week]"), with four of the items worded in the positive direction (Radloff, 1977). Questions refer to symptom presence in the last 4 weeks ("In the past month"). Because the CES-D-m cannot determine psychiatric diagnosis, MDD outcomes as determined by the CES-D-m were labeled "elevated MDD risk." Diagnostic assessment using the DIS module for MDD, administered by social workers trained by a DIS-certified psychiatrist, was used to assess full criteria for the diagnosis of MDD.

We used sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), overall correct classification (OCC), receiver operating characteristics (ROC), and Youden's index (J) to assess performance of the CES-D-m in the retiree population relative to the DIS diagnosis of MDD. Sensitivity measures a screening tool's ability to correctly identify those with a disorder. Specificity measures the tool's ability to correctly identify those without a disorder. PPV is the probability of having a disease when the tool tests positive and NPV is the probability of not having a disease when the tool tests negative. OCC is the percent of correctly identified participants with and without the disease. Youden's index (sensitivity + specificity − 1) allows for the selection of a cutoff score (the cutoff score at the maximum value of J) when sensitivity and specificity are equally important. A perfect screening tool relative to a diagnostic assessment would have a Youden's index of 1 (Bewick et al., 2004). We used the area underneath the ROC curve (AUC) to quantify the performance of the CES-D-m. An AUC of 1.00 indicates a perfect screening tool relative to a diagnostic assessment.

2.5. Statistical analyses

We analyzed demographic and other variables in relation to elevated MDD risk. Bivariate analyses of categorical variables used χ^2 with odds ratios (OR) and 95% confidence intervals (CI_{95}) or McNemar's test, as appropriate. Correlation between two variables was measured using Pearson's R^2. Continuous variables were assessed using ANOVA. Separate multivariate logistic regression analyses were performed to compare independent predictors of depression using three different outcomes: CES-D-m positive scored with two different thresholds, score 16 and score 22, and DSM-IV-TR® MDD diagnosis using the MDD section of the DIS. Backwards stepwise elimination was used to determine the final models, with variables retained if at least one of the categories was significant at $p \leq 0.10$. Goodness-of-fit was assessed using the Hosmer–Lemeshow test. All analyses were performed using Statistical Analysis Software, version 9.1.3 (SAS Institute, Cary, N.C.).

3. Results

3.1. Characteristics of the study population

As shown in Table 1, the 1915 retired firefighters averaged 47.0 years of age on 9/11 (SD = 6.9; range 28.0 to 73.4 years). The participants were predominantly married (83%) and about half had some college education or an associate degree (48%). At the time of the exam, 4 to 6 years after 9/11, 10% reported a change in marital status.

Table 1
Characteristics and average CES-D-m scores of retired firefighters.

	n	%	CES-D-m (mean score)
Total	1915	100	14.39
Age on 9/11 (*p*-value)			<0.0001
<40	292	15.25	15.35
40–44	470	24.54	14.37
45–49	563	29.40	15.50
50–54	323	16.87	14.46
55+	267	13.94	10.97
Exposure group (*p*-value)			<0.0001
1 — Morning of 9/11	297	15.51	17.32
2 — Afternoon of 9/11	1077	56.24	14.36
3 — Day 2	316	16.50	13.13
4 — Day 3 to day 14	225	11.75	12.45
Retirement status (*p*-value) [a]			<0.0001
Disability	1211	64.38	15.37
Non-disability	670	35.62	12.70
Previous professions (*p*-value) [b]			<0.01
No other profession	1030	53.79	15.03
Other professions	885	46.21	13.66
Number of children (*p*-value)			NS
No children (0)	195	10.18	14.46
1	218	11.38	13.87
2	735	38.38	14.30
3	510	26.63	14.39
4	183	9.56	15.46
5+	74	3.86	14.09
Education (*p*-value)			NS
High school diploma	545	28.46	14.71
College classes or associate degree	927	48.41	14.45
Bachelor degree or RN	309	16.14	14.15
Post college classes or degree	134	7.00	13.30
Rank (*p*-value)			<0.001
Chiefs	122	6.37	11.66
Captains and lieutenants	597	31.17	13.53
Firefighters	1196	62.45	15.10
Marital status on 9/11 (*p*-value) [c]			NS
Single	101	5.28	14.92
Married	1594	83.28	14.23
Not married, living with partner	63	3.29	15.30
Separated	34	1.78	16.15
Divorced	110	5.71	15.39
Widowed	12	0.63	13.58
Marital status change since 9/11 (*p*-value)			<0.05
No change	1726	90.13	14.21
Status change	189	9.87	16.03
Years retired (*p*-value) [d]			<0.05
Years ≤ 1	247	12.90	13.82
1<years≤2	282	14.73	14.73
2<years≤3	289	15.09	16.35
3<years≤4	614	32.06	14.11
4<years≤5	275	14.36	13.25
Years>5	208	10.86	14.25
Months spent at site (*p*-value) [e]			NS
Months<5	580	61.90	14.37
Months≥5	357	38.10	15.23
Years of FDNY service on 9/11 (*p*-value)			<0.0001
Years ≤ 15	403	21.04	15.97
15<years≤20	630	32.90	14.98
20<years≤25	546	28.51	14.42
Years>25	336	17.55	11.37

[a] Retirement pension information (disability vs. non-disability) was not available for those who retired before 1991 (n = 34).
[b] This referred specifically to previous professions of military, law enforcement, corrections, and health.
[c] One participant did not respond.
[d] Number of years the individual had been retired on the day the questionnaire was completed.
[e] The amount of time spent at the WTC site was available for only 937 (48.9%) of the participants.

Most participants arrived at the WTC site on 9/11 (72%), with the majority arriving after the collapse of the WTC towers, during the afternoon of 9/11 (56%). By the end of day 2, more than 88% reported working at the WTC site. Approximately 40% worked at the WTC site for 5 or more months.

Disability retirement pensions were granted to 64% of our study population. A significant association was observed between retirement with a disability pension and age ($p<0.0001$). When analyzed by the categories in Table 1, we found a 50% decrease in the probability of retirement on disability with every 10 year increase in age (OR = 0.50, CI_{95}: 0.50–0.51).

We compared characteristics of the 1915 participants to 2127 non-participating retirees who met the study criteria for inclusion. The primary difference between participants and non-participants was that the former lived in the tri-state area (NY, NJ and CT) ($p<0.01$). Participants were also older (47.0 vs. 45.4 years; $p<0.001$) and more likely to have retired as chiefs (6% vs. 4%; $p<0.0001$). Retiree groups were similar in the proportion of members who were present during the morning of the collapse (16% vs. 17% $p=0.21$), and in those retiring with disability benefits (65% vs. 64% $p=0.81$).

3.2. Elevated MDD risk using the CES-D-m and diagnosis of MDD using the DIS

The average CES-D-m score for retired firefighters was 14.4 (SD = 10.6). Age on 9/11, 9/11 exposure group, retirement status, previously held professions, rank, change in marital status since 9/11, number of years retired, number of months spent at the WTC site and years of FDNY service were all related to mean CES-D-m scores (Table 1). Using a CES-D-m cutoff score of 16, 36% of retired firefighters ($n=687$) were identified as having elevated MDD risk. In contrast, 7% of retired firefighters ($n=133$) were identified as having MDD using the DIS.

Table 2
Characteristics of the CES-D-m with different cutoff scores in relation to the Diagnostic Interview Schedule.

Cutoff score	Sensitivity	Specificity	PPV[a]	NPV[b]	OCC[c]	J[d]
10	0.970	0.430	0.113	0.995	0.467	0.400
11	0.962	0.475	0.120	0.994	0.509	0.438
12	0.962	0.525	0.131	0.995	0.555	0.487
13	0.955	0.590	0.148	0.994	0.616	0.545
14	0.955	0.624	0.159	0.995	0.647	0.579
15	0.955	0.656	0.172	0.995	0.677	0.611
16	**0.932**	**0.684**	**0.180**	**0.993**	**0.701**	**0.616**
17	0.917	0.711	0.192	0.991	0.725	0.628
18	0.910	0.736	0.204	0.991	0.748	0.645
19	0.895	0.758	0.216	0.990	0.768	0.653
20	0.872	0.778	0.227	0.988	0.784	0.650
21	0.865	0.797	0.242	0.987	0.802	0.662
22	**0.842**	**0.819**	**0.258**	**0.986**	**0.821**	**0.661**
23	0.812	0.837	0.271	0.984	0.836	0.649
24	0.782	0.855	0.287	0.981	0.850	0.637
25	0.752	0.877	0.313	0.979	0.868	0.629
26	0.707	0.891	0.325	0.976	0.878	0.597
27	0.662	0.898	0.327	0.973	0.882	0.560
28	0.632	0.911	0.347	0.971	0.892	0.543
29	0.602	0.920	0.360	0.969	0.898	0.522
30	0.556	0.931	0.376	0.966	0.905	0.487

[a]Positive predictive value.
[b]Negative predictive value.
[c]Overall correct classification.
[d]Youden's *J* index.
Bold indicates the current cutoff score of 16 being used at the FDNY and the optimal cutoff score of 22 using ROC analysis.

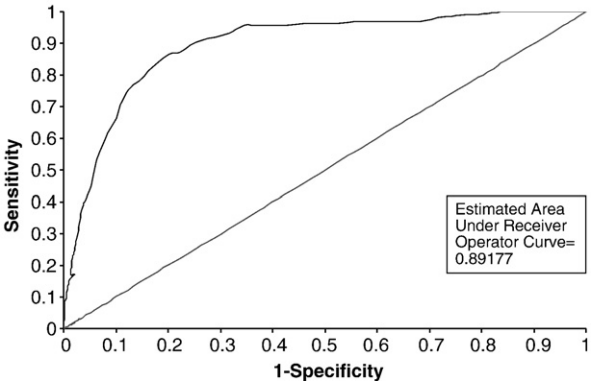

Fig. 1. ROC curve of CES-D-m taken by retired firefighters.

3.3. Diagnostic value

The ROC curve for the CES-D-m is presented in Fig. 1. The AUC was 0.89. We examined the diagnostic value of scores from 10 to 30 in comparison to the diagnosis of MDD determined to meet full DSM-IV-TR® symptom criteria by DIS interview (Table 2). A cutoff score of 16 had a sensitivity of 0.93 and a specificity of 0.68 in comparison with DIS diagnosis. However, ROC analysis and Youden's index were at their maximum for cutoff score 22 ($J=0.66$), which had a sensitivity of 0.84 and specificity of 0.82. In increasing the cutoff score from 16 to 22, elevated MDD risk was identified in only 23% of retired firefighters ($n=434$); the PPV increased by 8% to 26%, corresponding to a decreased in false positives from 563 to 322; and the NPV decreases by only 0.7%, corresponding to a increase in false negatives from 9 to 21 (Fig. 2).

3.4. Multivariate models

We used multivariate logistic regression models to estimate the independent effects of demographic and other variables on three outcomes: CES-D-m positive using separate cutoff scores of 16 and 22; and meeting full DSM-IV-TR® criteria for MDD using the DIS. We found that age on 9/11, exposure group, retirement disability status (ordinary retirement as reference), marital status change, and years of service to be statistically significant covariates in final models (Table 3). Elevated MDD risk was 1.42 to 1.52 times greater for those with less than or equal to 25 years of service. As shown in Model 1, exposure group, retirement status and years of FDNY service on 9/11 were statistically significant correlates of elevated MDD risk when a cutoff score of 16 was used. When the cutoff score was raised to 22 (Model 2), exposure group and retirement disability status remained statistically significant correlates, while years of FDNY service on 9/11 were no longer significant, but were instead replaced by age on 9/11 and martial status change since 9/11. When predicting MDD on the DIS, only age on 9/11 remained statistically significant (Model 3).

MDD was 2.12 to 5.35 times greater in younger age groups relative to the oldest retirees, and elevated MDD risk was 2.02

Table 3
Multivariate analysis of MDD models using logistic regression.

	Model 1		Model 2		Model 3	
	Elevated MDD risk with cutoff score 16 on CES-D-m		Elevated MDD risk with cutoff score 22 on CES-D-m		MDD on DIS module F	
Variable	Odds ratio	95% CI	Odds ratio	95% CI	Odds ratio	95% CI
Age on 9/11						
<40	_	_	2.47	1.49–4.10	5.35	2.20–13.02
40–44	_	_	2.17	1.34–3.51	3.29	1.36–7.95
45–49	_	_	2.96	1.85–4.73	3.96	1.67–9.39
50–54	_	_	2.31	1.39–3.83	2.12	0.81–5.54
55+	_	_	1.00	(ref)	1.00	(ref)
Exposure group						
1 — Morning of 9/11	2.02	1.37–2.98	2.55	1.60–4.10	_	_
2 — Afternoon of 9/11	1.26	0.90–1.76	1.47	0.96–2.26	_	_
3 — Day 2	1.04	0.70–1.54	1.27	0.78–2.07	_	_
4 — Day 3 to day 14	1.00	(ref)	1.00	(ref)	_	_
Retirement status						
Disability	1.55	1.25–1.92	1.79	1.40–2.31	_	_
Non-disability	1.00	(ref)	1.00	(ref)	_	_
Marital status change since 9/11 (p-value)						
Status change	_	_	1.41	1.00–1.99	_	_
No change	_	_	1.00	(ref)	_	_
Years of FDNY service on 9/11 (p-value)						
Years ≤ 15	1.42	0.98–2.06	_	_	_	_
15 < years ≤ 20	1.43	1.01–2.00	_	_	_	_
20 < years ≤ 25	1.52	1.09–2.11	_	_	_	_
Years > 25	1.00	(ref)	_	_	_	_

The following covariates were included in each model: age on 9/11, exposure group, retirement status, rank, years retired, months spent at site, change in marital status, and years of FDNY service on 9/11; variables presented were retained in final models.

to 2.55 times greater in retired firefighters who first arrived at the WTC site during the morning of 9/11. Persons who retired due to a disability were 1.55 to 1.79 times more likely to have elevated MDD risk as ordinary retirees.

3.5. Active firefighters

Active firefighters who took the mental health questionnaires between 12/2006 and 12/2007 ($n = 4543$) were investigated for comparison with retirees. Active firefighters were significantly younger, with an average age of 35.9 years on 9/11 (SD = 7.0; $p < 0.0001$). In addition, compared with retirees, active firefighters had significantly lower mean CES-D-m scores (10.0 vs. 14.4; SD = 8.2; $p < 0.0001$) and significantly lower prevalence of elevated MDD risk at cutoff scores of 16 and 22 ($p < 0.0001$). Compared to retirees, at CES-D-m cutoff scores of 16 and 22, 16% ($n = 736$) and 9% ($n = 392$) of active firefighters, respectively, were identified as having elevated MDD risk.

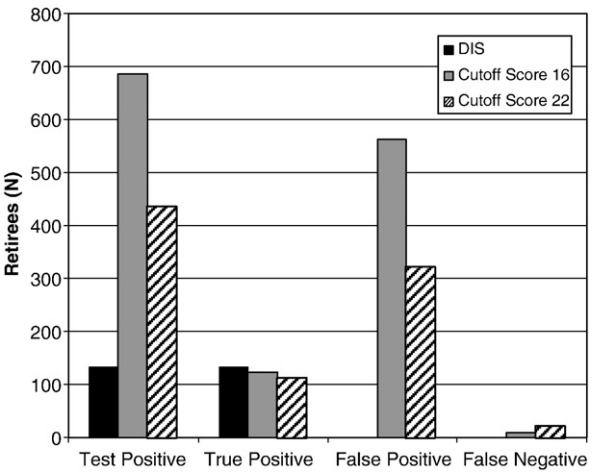

Fig. 2. Distribution of CES-D-m Cutoff Score Test results compared to the Diagnostic Interview Schedule.

4. Discussion

This study tests the performance of the CES-D-m in comparison to DIS identification of MDD in WTC-exposed retired FDNY firefighters. Previous studies in other populations have used the CES-D in its original form, assessing symptom occurrence in the past week (Beekman et al., 1995; Cloninger et al., 2006; Schein and Koenig, 1997; Zauszniewski and Graham, 2009). The DSM-IV-TR® criteria require symptoms to extend beyond the immediate 7 days prior to screening. We chose to modify the CES-D to assess a one-month duration of symptoms rather than two-week duration (DSM-IV-TR® definition), for consistency with the time frame of other questionnaires used in the FDNY WTC self-administered mental health monitoring exam (e.g. Posttraumatic stress disorder checklist) and to cast a wider net to identify those with elevated MDD risk. Overall, our results indicated that the CES-D-m with an AUC of 0.89 was a valuable instrument for identifying retired firefighters at elevated risk of depressive symptoms (Swets, 1988). Although the CES-D-m focused on symptoms over the past month, the AUC was comparable to findings of previous studies using the original CES-D, which focused on symptoms over the past week (Cuijpers et al., 2008; Ghubash et al., 2000; Haringsma et al., 2004; Jones et al., 2005; Watson et al., 2004). An earlier study, in a community based sample, similarly modified the CES-D time frame to 1 month and found that

although the modified instrument identified a slightly higher level of depressive symptoms compared to the standard CES-D, both instruments had the same high internal consistency ($\alpha = 0.89$) (Moscato et al., 1997).

This study underscores several basic tenets: that screening instruments when modified require validation in the population of interest and that the an optimal cutoff score for that population should be based on the prevalence of undiagnosed disease in the target population, the availability of interventions and resources, and the relative importance of identifying all true positives. As sensitivity and specificity vary by cutoff score, it can be difficult to determine the most appropriate cutoff score. While the cutoff score of 16 (Radloff, 1977) was acceptable, increasing the score to 22 optimized specificity with only a minor reduction in sensitivity.

Other validation studies in adolescents (Cuijpers et al., 2008) and the elderly (Cheng and Chan, 2005, 2008), populations that are demographically very different from FDNY firefighters, have also found a cutoff score of 22 to be optimal. Without validation, a cutoff score of 22 was used in a study of firefighters who responded to hurricane Katrina (Saijo et al., 2008) and NYC Transit Workers located near the WTC (Tapp et al., 2005). To our knowledge, no study, has validated the performance of the CES-D against a structured diagnostic interview (e.g. DIS) in rescue/recovery workers exposed to the WTC or other disasters.

WTC-exposed populations other than firefighters have been studied for depression but, those studies were carried out closer in time to 9/11 and used instruments other than the CES-D, such as the Patient Health Questionnaire (PHQ), State-Trait Personality Inventory-form Y, and Structured Clinical Interview for DSM-IV (SCID) (Galea et al., 2002; Gross et al., 2006; Long et al., 2007; Stellman et al., 2008). A study using the SCID conducted in 2001 on Manhattan residents living south of 110th Street (up to 6.75 miles north of the WTC) reported 9.7% with depression-like symptoms; this increased to 16.8% among those living less than 1 mile from the site (Galea et al., 2002). Twelve percent of a sample of NYC transit workers stationed near the WTC towers screened in 2002 using the CES-D (Tapp et al., 2005) and 8.8 to 16.1% of a sample of WTC clean up and recovery workers screened in 2003 using the PHQ reported symptoms consistent with major depression (Gross et al., 2006; Stellman et al., 2008). We found that 23% of retired firefighters reported symptoms consistent with major depression using the CES-D-m cutoff score of 22 and 7% had MDD using a full diagnostic assessment (DIS). Agreement between these screening instruments has not been tested in WTC or other disaster exposed populations and remains unclear in studies performed on chronically ill patients (Dbouk et al., 2008; Jones et al., 2005; Lowe et al., 2004). The PHQ and the CES-D were found to have excellent internal consistency compared to the SCID in outpatients and epileptics respectively (Jones et al., 2005; Lowe et al., 2004) but there was low agreement between the PHQ and CES-D in patients with chronic hepatitis C (Dbouk et al., 2008).

Our study population consisted of retirees with an average age of 47.0 years on 9/11. The youngest retirees had significantly higher average CES-D-m scores than older retirees. Disability pensions were also granted to a significantly higher percentage of younger retirees. At FDNY, age is positively correlated with years of service ($R^2 = 0.76$), as actives typically retire after 20 years of service when they qualify for regular pensions whereas those with work-related injuries/illnesses may qualify for disability pensions with fewer years of service. Those granted disability pensions, had significantly higher average CES-D-m scores (15.4) than those who did not retire on a disability pension (12.7). This difference in mean CES-D-m scores between retiree groups occurred despite exclusion of individuals who retired solely due to a mental health disability ($N = 31$) from all analyses in an effort to protect against inflating the relationship between disability retirement and MDD. This supports the idea that the mental health consequences of 9/11 may have been underappreciated in relation to physical illnesses, and suggests the need to further investigate co-morbid psychopathology and the relationship between MDD and physical disability. It also argues for programs that address the loss of social support when firefighters retire and leave the firehouse and work environment. Studies in other populations have found retirement, especially early retirement, to be associated with depression (Alavinia and Burdorf, 2008; Karpansalo et al., 2005) and social support to be protective (Pietrzak et al., 2009).

We also found a strong exposure-response gradient. Firefighters in exposure Group I (arrived morning of 9/11/01) were at the WTC site either during the collapse or shortly thereafter, generally suffered the greatest physical (Prezant et al., 2002) and psychological trauma (Fountoulakis et al., 2001) and had greater average CES-D-m scores (17.3) than exposure groups with later arrival times (12.5–14.4). This is consistent with previous findings in this and other WTC cohorts using PTSD screening instruments (Corrigan et al., in press; Galea et al., 2002). We minimized recall bias by using exposure information from the earliest post-9/11 assessments. In our multivariate models, exposure groups had a dose response relationship with MDD using either cutoff score of the CES-D-m. Because 9/11 exposure group is such an important factor in predicting MDD in this and other populations, it may be useful to consider adjustment of cutoff scores used in screening tools for those arriving earliest to a disaster to maximize identification of true positives.

A limitation of this study was that retiree participants were self-selected and were significantly more likely to live in the NYC area. Although we cannot rule out the possibility that this may have been associated with MDD since we do not have MDD data on non-participants, we are reassured that retiree participants were generally similar to retiree non-participants in their proportions of the most highly exposed and disabled persons. Participants were older and more likely to retire as officers compared with non-participants, both variables that were related to lower mean MDD scores in bivariate analyses. Although we believe it reasonable to generalize our CES-D-m threshold scores to our retiree population, it should be noted that our population was predominantly Caucasian males and therefore validations studies should be performed before generalizing to other populations.

5. Conclusions

We believe that the CES-D-m is an effective way to screen for MDD in retired male firefighters. On the basis of ROC analyses and Youden's index, where neither sensitivity nor specificity is favored, the optimal cutoff score in our

population was 22. This cutoff score is higher than what was estimated by the original study carried out in the general population rather than in firefighters or other disaster workers (Radloff, 1977). Studies in other populations, disaster and non-disaster exposed, have also found a cutoff score of 16 to be too low (Cheng and Chan, 2005; Cheng and Chan, 2008; Cloninger et al., 2006; Cuijpers et al., 2008; Fountoulakis et al., 2001; Goebert et al., 2009; Haringsma et al., 2004; Tak et al., 2007; Tapp et al., 2005; Wada et al., 2007). Identifying MDD in retired firefighters allows for prompt referral for complete assessment and appropriate consideration of treatment, which may reduce impairment of the affected individual. Further work is necessary to establish an optimal cutoff score for active firefighters where MDD not only influences the affected individual, but also may negatively impact job performance during critical firefighting and rescue work.

Role of funding source

Funding for this work was provided by the National Institute for Occupational Safety and Health (NIOSH) RO1-OH07350. NIOSH had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgements

This work was supported by the National Institute for Occupational Safety and Health.

References

Alavinia, S.M., Burdorf, A., 2008. Unemployment and retirement and ill-health: a cross-sectional analysis across European countries. Int. Arch. Occup. Environ. Health 82 (1), 39–45.

Beekman, A.T., Deeg, D.J., van, T.T., Smit, J.H., Hooijer, C., van, T.W., 1995. Major and minor depression in later life: a study of prevalence and risk factors. J. Affect. Disord. 36 (1–2), 65–75.

Bewick, V., Cheek, L., Ball, J., 2004. Statistics review 13: receiver operating characteristic curves. Crit. Care 8 (6), 508–512.

Cheng, S.T., Chan, A.C., 2005. The Center for Epidemiologic Studies Depression Scale in older Chinese: thresholds for long and short forms. Int. J. Geriatr. Psychiatry 20 (5), 465–470.

Cheng, S.T., Chan, A.C., 2008. Detecting depression in Chinese adults with mild dementia: findings with two versions of the Center for Epidemiologic Studies Depression Scale. Psychiatry Res. 159 (1–2), 44–49.

Cloninger, C.R., Svrakic, D.M., Przybeck, T.R., 2006. Can personality assessment predict future depression? A twelve-month follow-up of 631 subjects. J. Affect. Disord. 92 (1), 35–44.

Corrigan, M., McWilliams, R., Kelly, K.J., Niles, J., Cammarata, C., Jones, K., Glass, L., Wartenberg, D., Halliman, W., Kipien, H., Schorr, J., Prezant, D.J., in press. Mental Health Outcomes In The FDNY Firefighter World Trade Center Cohort —A Study Based On A Self-Administered, Computerized Survey With 2.5 Years Of Follow Up. Am. J. Public Health.

Cuijpers, P., Boluijt, P., van Straten, A., 2008. Screening of depression in adolescents through the Internet: sensitivity and specificity of two screening questionnaires. Eur. Child Adolesc. Psychiatry 17 (1), 32–38.

Dbouk, N., Arguedas, M.R., Sheikh, A., 2008. Assessment of the PHQ-9 as a screening tool for depression in patients with chronic hepatitis C. Dig. Dis. Sci. 53 (4), 1100–1106.

Fountoulakis, K., Iacovides, A., Kleanthous, S., Samolis, S., Kaprinis, S.G., Sitzoglou, K., St, K.G., Bech, P., 2001. Reliability, validity and psychometric properties of the Greek translation of the Center for Epidemiological Studies—Depression (CES-D) Scale. BMC Psychiatry 1, 3.

Galea, S., Ahern, J., Resnick, H., Kilpatrick, D., Bucuvalas, M., Gold, J., Vlahov, D., 2002. Psychological sequelae of the September 11 terrorist attacks in New York City. N. Engl. J. Med. 346 (13), 982–987.

Ghubash, R., Daradkeh, T.K., Al Naseri, K.S., Al Bloushi, N.B., Al Daheri, A.M., 2000. The performance of the Center for Epidemiologic Study Depression Scale (CES-D) in an Arab female community. Int. J. Soc. Psychiatry 46 (4), 241–249.

Goebert, D., Thompson, D., Takeshita, J., Beach, C., Bryson, P., Ephgrave, K., Kent, A., Kunkel, M., Schechter, J., Tate, J., 2009. Depressive symptoms in medical students and residents: a multischool study. Acad. Med. 84 (2), 236–241.

Gross, R., Neria, Y., Tao, X.G., Massa, J., Ashwell, L., Davis, K., Geyh, A., 2006. Posttraumatic stress disorder and other psychological sequelae among World Trade Center clean up and recovery workers. Ann. N. Y. Acad. Sci. 1071, 495–499.

Haringsma, R., Engels, G.I., Beekman, A.T., Spinhoven, P., 2004. The criterion validity of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of self-referred elders with depressive symptomatology. Int. J. Geriatr. Psychiatry 19 (6), 558–563.

Jones, J.E., Hermann, B.P., Woodard, J.L., Barry, J.J., Gilliam, F., Kanner, A.M., Meador, K.J., 2005. Screening for major depression in epilepsy with common self-report depression inventories. Epilepsia 46 (5), 731–735.

Karpansalo, M., Kauhanen, J., Lakka, T.A., Manninen, P., Kaplan, G.A., Salonen, J.T., 2005. Depression and early retirement: prospective population based study in middle aged men. J. Epidemiol. Community Health 59 (1), 70–74.

Katz, C.L., Pellegrino, L., Pandya, A., Ng, A., DeLisi, L.E., 2002. Research on psychiatric outcomes and interventions subsequent to disasters: a review of the literature. Psychiatry Res. 110 (3), 201–217.

Long, M.E., Meyer, D.L., Jacobs, G.A., 2007. Psychological distress among American Red Cross disaster workers responding to the terrorist attacks of September 11, 2001. Psychiatry Res. 149 (1–3), 303–308.

Lowe, B., Spitzer, R.L., Grafe, K., Kroenke, K., Quenter, A., Zipfel, S., Buchholz, C., Witte, S., Herzog, W., 2004. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. J. Affect. Disord. 78 (2), 131–140.

Mair, C., Diez Roux, A.V., Shen, M., Shea, S., Seeman, T., Echeverria, S., O'Meara, E.S., 2009. Cross-sectional and longitudinal associations of neighborhood cohesion and stressors with depressive symptoms in the multiethnic study of atherosclerosis. Ann. Epidemiol. 19 (1), 49–57.

Moscato, B.S., Russell, M., Zielezny, M., Bromet, E., Egri, G., Mudar, P., Marshall, J.R., 1997. Gender differences in the relation between depressive symptoms and alcohol problems: a longitudinal perspective. Am. J. Epidemiol. 146 (11), 966–974.

North, C.S., Nixon, S.J., Shariat, S., Mallonee, S., McMillen, J.C., Spitznagel, E.L., Smith, E.M., 1999. Psychiatric disorders among survivors of the Oklahoma City bombing. JAMA 282 (8), 755–762.

Person, C., Tracy, M., Galea, S., 2006. Risk factors for depression after a disaster. J. Nerv. Ment. Dis. 194 (9), 659–666.

Pietrzak, R.H., Johnson, D.C., Goldstein, M.B., Malley, J.C., Southwick, S.M., 2009. Psychological resilience and postdeployment social support protect against traumatic stress and depressive symptoms in soldiers returning from Operations Enduring Freedom and Iraqi Freedom. Depress. Anxiety. 26 (8), 745–747.

Prezant, D.J., Weiden, M., Banauch, G.I., McGuinness, G., Rom, W.N., Aldrich, T.K., Kelly, K.J., 2002. Cough and bronchial responsiveness in firefighters at the World Trade Center site. N. Engl. J. Med. 347 (11), 806–815.

Radloff, L.S., 1977. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Appl. Psychol. Meas. 1 (3), 385–401.

Robins, L.N., Cottler, L., Bucholz, K., Compton, W., 1995. Diagnostic Interview Schedule for DSM-IV. Washington University, St Louis.

Saijo, Y., Ueno, T., Hashimoto, Y., 2008. Twenty-four-hour shift work, depressive symptoms, and job dissatisfaction among Japanese firefighters. Am. J. Ind. Med. 51 (5), 380–391.

Schein, R.L., Koenig, H.G., 1997. The Center for Epidemiological Studies—Depression (CES-D) Scale: assessment of depression in the medically ill elderly. Int. J. Geriatr. Psychiatry 12 (4), 436–446.

Schiepers, O.J., de Groot, R.H., Jolles, J., van Boxtel, M.P., 2009. Plasma phospholipid fatty acid status and depressive symptoms: association only present in the clinical range. J. Affect. Disord. 118 (1-3), 209–214.

Shinar, D., Gross, C.R., Price, T.R., Banko, M., Bolduc, P.L., Robinson, R.G., 1986. Screening for depression in stroke patients: the reliability and validity of the Center for Epidemiologic Studies Depression Scale. Stroke 17 (2), 241–245.

Snoek, F.J., van,d., V., Twisk, J.W., Hogenelst, M.H., Tromp-Wever, A.M., van der Ploeg, H.M., Heine, R.J., 2008. Cognitive behavioural therapy (CBT) compared with blood glucose awareness training (BGAT) in poorly controlled Type 1 diabetic patients: long-term effects on HbA moderated by depression. A randomized controlled trial. Diabet. Med. 25 (11), 1337–1342.

Stellman, J.M., Smith, R.P., Katz, C.L., Sharma, V., Charney, D.S., Herbert, R., Moline, J., Luft, B.J., Markowitz, S., Udasin, I., Harrison, D., Baron, S., Landrigan, P.J., Levin, S.M., Southwick, S., 2008. Enduring mental health morbidity and social function impairment in World Trade Center rescue, recovery, and cleanup workers: the psychological dimension

of an environmental health disaster. Environ. Health Perspect. 116 (9), 1248–1253.
Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240 (4857), 1285–1293.
Tak, S., Driscoll, R., Bernard, B., West, C., 2007. Depressive symptoms among firefighters and related factors after the response to Hurricane Katrina. J. Urban Health 84 (2), 153–161.
Tapp, L.C., Baron, S., Bernard, B., Driscoll, R., Mueller, C., Wallingford, K., 2005. Physical and mental health symptoms among NYC transit workers seven and one-half months after the WTC attacks. Am. J. Ind. Med. 47 (6), 475–483.
Wada, K., Tanaka, K., Theriault, G., Satoh, T., Mimura, M., Miyaoka, H., Aizawa, Y., 2007. Validity of the Center for Epidemiologic Studies Depression Scale as a screening instrument of major depressive disorder among Japanese workers. Am. J. Ind. Med. 50 (1), 8–12.
Watson, L.C., Lewis, C.L., Kistler, C.E., Amick, H.R., Boustani, M., 2004. Can we trust depression screening instruments in healthy 'old–old' adults? Int. J. Geriatr. Psychiatry 19 (3), 278–285.
Zauszniewski, J.A., Graham, G.C., 2009. Comparison of short scales to measure depressive symptoms in elders with diabetes. West. J. Nursing Res. 31 (2), 219–234.