

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/292211195>

Effective Screening for Emotional Distress in Refugees

Article *in* Journal of Nervous & Mental Disease · January 2016

Impact Factor: 1.69 · DOI: 10.1097/NMD.0000000000000469

READS

37

7 authors, including:



Michael Hollifield

VA Long Beach Healthcare System

65 PUBLICATIONS 2,779 CITATIONS

[SEE PROFILE](#)



Beth Farmer

Department of Family and Community Ser...

2 PUBLICATIONS 9 CITATIONS

[SEE PROFILE](#)

Effective Screening for Emotional Distress in Refugees

The Refugee Health Screener

Michael Hollifield, MD,*† Eric C. Toolson, PhD,‡ Sasha Verbillis-Kolp, MSW,§ Beth Farmer, LICSW,§ Junko Yamazaki, LICSW,|| Tsegaba Woldehaimanot, MSW,|| and Annette Holland, BS¶

Abstract: Screening for emotional distress is important, but not widely available. This study assesses the utility of the Refugee Health Screener 15 (RHS-15) in a public health setting. Refugee Health Screener 15 and diagnostic proxy (DP) instruments assessing anxiety, depression, and posttraumatic stress disorder were administered to refugees from 3 countries at their public health examination. Properties of the RHS-15 and its components were evaluated utilizing appropriate methods. Scale Cronbach α was 0.95, and a factor analysis identified 1 factor accounting for 66% of scale variance. Refugee Health Screener 15 scores and cases discriminated between refugee groups similar to DPs. Refugee Health Screener 15 case sensitivity and specificity to DPs were acceptable ($\geq 0.87/0.77$). A shorter, 13-item component had acceptable metric properties. The RHS-15 appears to be a valid screener for emotional distress of refugees. The 13-item scale may be more efficient and as efficacious for case identification. The critical public health need and recommendations for implementation are discussed.

Key Words: Distress, mental health, psychometric, refugees, RHS-15, screening (*J Nerv Ment Dis* 2016;00: 00–00)

There are now approximately 59.5 million people (40 million with refugee status) worldwide who are forcibly displaced from their homes because of war, violence, and oppression, a rapidly changing increase in the past two years due to the Syrian crisis. Eighty-six percent of refugees are hosted in developing countries (United Nations High Commissioner for Refugees, 2015). From 1975 to 2009, more than 1.8 refugees resettled in the United States (Services, Bridging Refugee Youth and Children's Services, September 2011; US Committee for Refugees and Immigrants World Refugee Survey, 2009), and with the current ceiling of 70,000 per year, there have now been more than 2 million refugees settled in the United States. A significant minority of refugees have diagnostic-level psychiatric disorders (Fazel et al., 2005; Fenta et al., 2004; Hollifield et al., 2009; Jaranson et al., 2004; Laban et al., 2004; Marshall et al., 2005; Momartin et al., 2006; Porter and Haslam, 2005; Weine et al., 1998), which are associated with stressful events in a dose-dependent manner (de Jong et al., 2001; Hollifield et al., 2006; Marshall et al., 2005; Mollica et al., 1998). The Office of Refugee Resettlement guidelines suggest a health screening of refugees in the first 90 days after arrival; however, despite support for mental health screening by the Centers for Disease Control (CDC, 2015), there has been a lack of procedural or financial support for

such screening (Savin et al., 2005). State refugee health coordinators surveyed in 2010 reported that only 4 of the 44 states surveyed used a formal screening instrument, and 68% used informal conversation (Shannon et al., 2012).

Poor screening has been in part due to the lack of an efficient or valid screener for common mental disorders in refugees and in part due to conceptual debates about what constitutes distress or "disorder" in refugees. The Refugee Health Screener 15 (RHS-15) was developed with these issues in mind. As previously reported, items were selected for the RHS-15 from 3 sources using 3 methods, relying heavily on a naive Bayesian classification procedure (Hollifield et al., 2013). Initial testing provided evidence for validity of the RHS-15, yet there are no reports about the instrument in its current form or analyses about what might be the most efficient set of items for screening. It is important to provide this information because there is significant interest in using the RHS: 167 US healthcare sites and 7 international sites have requested it for clinical assessment or evaluation. This is the first report of the properties of the RHS-15 administered in a public health setting, in this case at the civil surgeon examination (8–12 months after the first domestic medical examination) at Public Health Seattle and King County (public health). Ethical review and approvals were conducted by the Pacific Institute for Research and Evaluation and the ethics committee at public health.

METHODS

Design and Hypotheses

Development and testing of the RHS-15 were part of the mission of the *Pathway to Wellness: Integrating Refugee Health and Wellbeing* (P2W) project. A cross-sectional administration of the RHS-15 and diagnostic proxies (DPs) for anxiety, depression, and posttraumatic stress disorder (PTSD) was conducted. Diagnostic proxies were used for efficiency; diagnostic interviews were not acceptable because of administration time and cost. The a priori hypothesis was that the RHS-15 would demonstrate good internal reliability ($\alpha > 0.80$) and validity (sensitivity and specificity both ≥ 0.85 to all DPs, discriminant validity between positive and negative RHS cases). Evaluation of the full RHS-15 and alternative RHS-15 components was conducted to determine if the RHS-15 is best utilized in its current form or shortened for efficiency.

Sample Frame, Sampling, and Procedures

The sample frame was a cohort of refugees 14 years or older from 3 countries (Bhutan, Burma, and Iraq) speaking 4 languages (Nepali, Karen, Burmese [Karenni and Chin ethnic groups], and Arabic) who were resettled in King County, Wash, and seen at public health for their health evaluation. Convention hereafter, and with respect to country, language, and ethnic group, will be to refer to participants by country of origin ("country": ie, Bhutanese, Burmese, and Iraqi). This sample frame was chosen because these were the most numerous refugees being resettled during the study period. Consecutive sampling of all eligible and available subjects ($N = 251$) was conducted on prespecified days by the P2W coordinator at public health during the domestic medical examination. Written informed consent for

*Pacific Institute for Research and Evaluation, Albuquerque, NM; †The VA Long Beach Healthcare System, Long Beach, CA; ‡Department of Biology, University of New Mexico, Albuquerque, NM; §Lutheran Community Services Northwest, SeaTac; ||Asian Counseling, Referral Service; and ¶Public Health Seattle & King County, Seattle, WA.

Project partners: Pathways to Wellness: Integrating Community Health and Well-being is a project of Lutheran Community Services Northwest, Asian Counseling and Referral Service, Public Health Seattle & King County, and Michael Hollifield, MD.

Send reprint requests to Michael Hollifield, MD, VA Long Beach Healthcare System, 5901 East 7th St, Long Beach, CA 90822. E-mail: michael.hollifield@va.gov.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 0022-3018/16/0000-0000

DOI: 10.1097/NMD.0000000000000469

all subjects was obtained. Subjects who returned to the civil surgeon examination ($n = 179$) were administered the RHS-15 by a trained staff nurse at public health, where the RHS-15 is now routinely utilized. The P2W coordinator administered the DPs within 2 weeks of the civil surgeon examination, blind to RHS-15 results.

Instruments and Data Analyses

Translation of Instruments

Translation is complex and must be adapted for specific purposes (Westermeyer and Janca, 1997). All instruments were translated using a rigorous, iterative back-and-forth participatory consensus process with refugees from each language group. This process ensured relevant language-specific semantics and cultural equivalence yielding accuracy and clarity of meaning across groups (Brislin, 1970; Hollifield et al., 2006).

The Refugee Health Screener 15

The RHS-15 has 13 symptom items, 1 coping item, and 1 distress thermometer (DT). The instructions for the symptom items are to “indicate the degree to which the symptom has been bothersome to you over the past month.” Possible responses are 0 = not at all, 1 = a little bit, 2 = moderately, 3 = quite a bit, and 4 = extremely. To help address variable literacy and cultural norms for understanding scales, there are symbols of jars with beans in them over the possible responses, with variable amounts of beans relevant to each response. The coping item assesses beliefs about general ability to cope with stress, and responses range from 0 (able to handle or cope with anything that comes your way) to 4 (unable to handle or cope with anything). The DT looks like a thermometer, with a “0” (“no distress—things are good”) at the bottom and a “10” (extreme distress—I feel as bad as I ever have”) at the top. The 13 symptom and 1 coping item responses are added to obtain a 14-item “total score.” Current recommended scoring is that an “RHS-15 case” is defined as a total score of 12 or greater or a DT 5 or greater, which is supported by previous post hoc testing (Hollifield et al., 2013).

Field testing and feedback from partner sites indicated that the coping item is the most time consuming and difficult for many refugees to understand, and the added value of the DT has been questioned. While the RHS-15 is efficient and accepted by the field, it is reasonable given the aforementioned feedback to evaluate the metric properties of a version without the coping item and/or the DT. Refugee Health Screener 15 components evaluated were (1) total score (continuous score adding 13 symptom and 1 coping items), (2) total case (dichotomous cutoff score of total score where positive case ≥ 12), (3) RHS case (total score of ≥ 12 or a DT ≥ 5), (4) DT score (continuous; range, 0–10), (5) DT case (≥ 5 DT score), (6) 13-item score (continuous score adding 13 symptom items), and (7) 13-item case (dichotomous cutoff score of 13-item score where positive case ≥ 11).

Diagnostic Proxy Instruments

Few instruments that assess symptoms as DPs in refugees are available (Hollifield et al., 2002). None are definitive diagnostic equivalents. The Hopkins Symptom Checklist 25 is a valid indicator of anxiety and depression for the general US population and for Indochinese refugees (Derogatis et al., 1974; Hollifield et al., 2002; Mollica et al., 1987; Winokur et al., 1984) and demonstrates transcultural validity (Butcher, 1991; Kinzie and Manson, 1987). Item-average scores of 1.75 or greater predict clinically significant anxiety (ANX) and depression (DEP) on the respective scales in general US and refugee samples and are considered valid DPs (Derogatis et al., 1974; Winokur et al., 1984).

The Posttraumatic Symptom Scale—Self report (PSS-SR) predicts *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* PTSD diagnosis in US populations (Foa et al., 1993). Cronbach α is 0.91, and 1-month test-retest reliability is 0.74. The 17 items on

the scale, each scored from 0 to 3 for symptom frequency, are *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* PTSD diagnostic items. The PSS-SR may be scored as continuous or a dichotomous DP. The PSS-SR continuous scores and the DP are highly correlated with war-related trauma and impairment in Kurdish and Vietnamese refugees (Hollifield et al., 2006), and Cronbach α in these samples was 0.95.

Data Management

The RHS-15 and DP data were entered into Excel spreadsheets by the P2W coordinator. Random spot checking of data accuracy on approximately 10% of cases was conducted by the coordinator. Two authors (M.H. and E.C.T.) evaluated data and excluded cases by each assessment if there were missing data of 10% or greater on an assessment. Valid cases were 178 RHS-15, total score, DT, and 13-item score, 175 PTSD and ANX, and 174 DEP, respectively. Interpolation of missing items on the RHS-15 for valid cases ($n = 8$) was conducted by using the mean score of completed items. Final Excel data were transferred to SPSS (IBM, version 18, Armonk, NY) for analyses.

Data Analyses

Descriptive data about the sample and the RHS-15 are reported. Generalized linear models were used to assess main and interactive effects on total scores, RHS cases, and 13-item scores and cases by country and sex. Because sex had no or modest effects on score and case determination, remaining analyses used country with combined sex samples. χ^2 Tests and formulas for positive predictive validity (PPV), negative predictive validity (NPV), and sensitivity and specificity were used to evaluate predictive validity of total scores, RHS cases, and 13-item cases. Concurrent validity was evaluated by description of ANX, DEP, and PTSD by country and correlations between these DPs and all RHS-15 components. Item-item correlations, internal scale reliability (Cronbach α), and principal component factor analyses set for significance with an eigenvalue (EV) of 1 or greater for the full sample are presented.

RESULTS

Descriptive

Sample

The cohort of 251 refugees 14 years or older was sampled at the domestic medical examination between April 2010 and November 2010 (Hollifield et al., 2013). One hundred seventy-nine (response rate, 71%) returned for and were evaluated at the civil surgeon examination (61 Iraqi, 55 Bhutanese, and 63 Burmese [31 Karen and 32 Burmese speaking]) between January 2011 and August 2011. One RHS-15 case was excluded because of missing data. The mean age of the sample was 32.4 ± 12.5 years; 93 were men, and 86 were women with no statistical differences between sex groups on age. The age of this sample (32.4 ± 12.5 years) was similar to that of the full cohort (31.8 ± 11.9 years). Female sex of this sample (48%) and the full cohort (46.7%) were similar.

RHS-15 Scores and Cases

Table 1 shows demographics and sample descriptions of the RHS-15 components. Overall group differences were primarily due to lower rates of distress for Burmese (8 positive cases, 13%) compared with Iraqi (37 positive cases, 61%) and Bhutanese (36 positive cases, 65%). Post hoc multiple comparisons showed a difference between Iraqi and Bhutanese on total score ($p = 0.13$) but not on DT score or RHS-15 case. Generalized linear model analysis for total score showed main effects for country (Wald = 25.1 [$df = 2$], $p < 0.001$) and interaction effects for country and sex (Wald = 9.1 [$df = 1$], $p = 0.003$) but no main effects for sex (Wald = 2.3 [$df = 1$], $p = 0.13$). Generalized linear model analysis for RHS-15 case showed main effects for country

TABLE 1. Comparison of Demographics and RHS-15 Components by Country of Origin, Score Means (SD), and Case Number (%) in Groups

	Iraq (n = 60)	Burma (n = 63)	Bhutan (n = 55)	Statistic	p
Age, y	32.2 (10.9)	33.3 (14.1)	31.5 (12.4)	F = 0.298	0.743
Female, n (%)	32 (52.4)	29 (46.0)	26 (47.3)	F = 0.281	0.756
Total score	21.7 (16.6)	5.9 (5.2)	15.6 (14.9)	F = 23.0	<0.001
DT score	4.9 (3.4)	1.2 (1.4)	4.3 (3.0)	F = 33.0	<0.001
Total case	35 (58%)	8 (13%)	27 (49%)	$\chi^2 = 30.0$	<0.001
DT case	31 (53%)	2 (3%)	28 (53%)	$\chi^2 = 42.2$	<0.001
RHS-15 case	37 (61%)	8 (13%)	36 (66%)	$\chi^2 = 41.3$	<0.001
13-Item score	20.3 (15.8)	4.8 (4.8)	13.6 (14.2)	F = 24.2	<0.001
13-Item case	35 (58%)	7 (11%)	25 (45%)	$\chi^2 = 31.3$	<0.001

Sex data are not included because of lack of main effects for sex.

RHS-15 components: “Total score” is sum of RHS-15 items 1–14; “Total case” is positive if total score is ≥ 12 ; “DT score” is the score on the DT, range 0–10; “DT case” is DT score ≥ 5 ; “RHS-15 case” is positive if either “total case” or “DT case” is positive; “13-item score” is sum of RHS-15 items 1–13; “13-item case” is positive if 13-item score is ≥ 11 .

(Wald = 33.8 [$df = 2$], $p < 0.001$) but not for sex (Wald = 3.5 [$df = 1$], $p = 0.06$) or for interaction between country and sex. Similar group differences were found for the shorter 13-item score and case criterion: post hoc multiple comparisons showed differences between all groups on score and significant differences between groups on case except for Iraqi-Bhutanese comparison ($p = 0.12$). Across groups, females tended to have higher 13-item scores than did males (mean, 14.8 [SD, 15.1] vs 10.8 [SD, 12.6]; $F = 3.9$; $p = 0.05$), but sex by 13-item case number was not significantly different (female: 36 cases [42%], male: 31 cases [34%]; $\chi^2 = 1.3$, $p = 0.26$). Generalized linear model analysis for 13-item case showed main effects for country (Wald = 25.6 [$df = 2$], $p < 0.001$), but not sex (Wald = 0.96 [$df = 1$], $p = 0.33$), and interaction effects for country by sex (Wald = 12.1 [$df = 1$], $p = 0.001$).

Reliability

For combined data, item correlations ranged from 0.360 to 0.802. Item 14, the coping item, had the lowest correlation with every other item, ranging from 0.360 to 0.510. Cronbach α 's for the full scale, total score, and 13-item score were 0.951, 0.952, and 0.960, respectively. Iraqi data showed $\alpha = 0.96$ (r range, 0.44–0.90) and 0.93 (r range, 0.45–0.90) on the full scale and 13-item score, respectively. Nepali data showed $\alpha = 0.94$ (r range, 0.32–0.80) and 0.96 (r range, 0.42–0.84), and Burmese data showed $\alpha = 0.82$ (r range, 0.13–0.65) and 0.80 (r range, -0.11 –0.62) on the respective component scales.

Factor Analysis

Analysis using all 15 items yielded an initial solution with 1 significant factor accounting for 66.3% of the variance, EV 9.9, communalities ranging from 0.30 (item 14) to 0.80 (item 8), and the component matrix values ranging from 0.55 to 0.89. Using total score items only (no DT) yielded 1 significant factor accounting for 66.5% of the variance, EV 9.3, communalities ranging from 0.31 (item 14) to 0.81 (item 8), and component matrix values ranging from 0.56 to 0.90. The 13 items only (no coping or DT items) yielded 1 significant factor accounting for 69.4% of the variance, EV 9.0, communalities ranging from 0.58 (item 1) to 0.81 (item 8), and component matrix values ranging from 0.76 to 0.90.

Multiple group analyses of the both 15- and 13-item components showed group differences. The 13 items yielded 1 component for Iraqi data (EV 9.3; variance 71.6%, communality range 0.58–0.87), 1 or 2 components for Nepali data (EV 8.7 and 1.0; variance 66.8% and 7.8%, communality range 0.61–0.85), and 4 possible components for Burmese data (EV 4.1, 1.8, 1.4, 1.2; variance 31.2, 13.9, 10.8, 9.3; communality range 0.47–0.78). In the case of the Burmese, items did not fit clearly in unrotated or rotated factors. All 15 items yielded similar data by country.

Validity

Concurrent Validity: DP Scores and Cases and Correlation With the RHS-15

Table 2 shows scores and case number by country for ANX, DEP, and PTSD. Data generally parallel RHS-15 findings, with Iraqi and Bhutanese groups having similar case rates (ANX 48% vs 39%;

TABLE 2. Comparison of Diagnostic Proxy Scores, Mean (SD), and Cases, n (%), by Country of Origin

	Iraq (n = 60)	Burma (n = 63)	Bhutan (n = 55)	Statistics
ANX score	1.91 (0.81)	1.17 (0.21)	1.68 (0.79)	Model F = 21.1; $p < 0.001$ Iraq-Burma $t = 7.1$; $p \leq 0.001$ Iraq-Bhutan $t = 1.7$; $p = 0.102$ Burma-Bhutan $t = 4.9$; $p \leq 0.001$
ANX case	28 (48)	1 (2)	21 (39)	$\chi^2 = 36.7$; $p < 0.001$
DEP score	2.05 (0.92)	1.20 (0.20)	1.76 (0.69)	Model F = 25.7; $p < 0.001$ Iraq-Burma $t = 7.2$; $p \leq 0.001$ Iraq-Bhutan $t = 1.9$; $p = 0.063$ Burma-Bhutan $t = 6.1$; $p \leq 0.001$
DEP case	29 (49)	1 (2)	26 (49)	$\chi^2 = 41.3$; $p < 0.001$
PTSD scores	18.3 (15.7)	3.4 (4.0)	14.9 (14.7)	Model F = 24.7; $p < 0.001$ Iraq-Burma $t = 7.4$; $p \leq 0.001$ Iraq-Bhutan $t = 1.3$; $p = 0.202$ Burma-Bhutan $t = 5.9$; $p \leq 0.001$
PTSD case	37 (63)	6 (10)	24 (45)	$\chi^2 = 38.3$; $p < 0.001$

Sex data are not included because of lack of main effects for sex.

Diagnostic proxy definitions: ANX score: the item-average on the 10 HSCL-25 (Hopkins Symptom Checklist 25) anxiety items; DEP score: the item-average on the 15 HSCL-25 depression items; PTSD score: the sum of all 17 items on the PSS-SR; ANX case: ≥ 1.75 item-average on the 10 HSCL-25 anxiety items; DEP case: ≥ 1.75 item-average on the 15 HSCL-25 depression items; PTSD case: meeting PTSD case definition of the PSS-SR.

DEP 49% vs 49%; PTSD 63% vs 45%), Iraqi having higher symptom scores than Bhutanese (ANX 1.91 [0.81] vs.1.68 [0.79]; DEP 2.05 [0.92] vs 1.71 [0.66]; PTSD 18.3 [15.7] vs 14.9 [14.7]), and Burmese with relatively low case rates and symptom scores.

Correlation between RHS-15 total score and ANX, DEP, and PTSD scores (and DPs) was 0.88 (0.83), 0.88 (0.81), and 0.91 (0.75), respectively. Table 3 shows correlations by group, which shows robust correlations in the Iraqi subsample, less robust in the Bhutanese group, and modest correlations in the Burmese group, likely due to low case rates.

Predictive Validity: Case Discrimination, PPV and NPV, and Sensitivity and Specificity

Table 4 shows that the total case, RHS case, and 13-item case all predict ANX, DEP, and PTSD DPs with high fidelity. Compared with the total case and the 13-item case, the RHS case has reduced specificity and PPV because of more false positives generated by the DT, whereas sensitivity of the RHS case is similar to total case and 13-item case on ANX, similar to total case on PTSD (slightly better than 13-item case), and slightly better than total case and 13-item case on DEP. Total case and 13-item case show similar characteristics, with slightly higher PPV and specificity for 13-item score case on ANX and DEP but not PTSD and slightly higher (but likely clinically insignificant except for PTSD) NPV and sensitivity for total case on all DPs.

TABLE 3. Correlations Between RHS-15 Components and Diagnostic Proxies by Country of Origin

	RHS-15 Component	ANX Score	ANX Case	DEP Score	DEP Case	PTSD Score	PTSD Case
Full sample	Total score	0.88	0.83	0.88	0.81	0.91	0.75
	Total case	0.74	0.77	0.76	0.76	0.80	0.77
	DT score	0.72	0.71	0.74	0.70	0.72	0.61
	DT case	0.64	0.65	0.68	0.67	0.63	0.54
	RHS case	0.68	0.69	0.72	0.72	0.73	0.68
	13-Item score	0.89	0.84	0.88	0.81	0.91	0.75
	13-Item case	0.75	0.77	0.77	0.76	0.78	0.73
Iraq	Total score	0.91	0.90	0.88	0.89	0.88	0.67
	Total case	0.77	0.84	0.79	0.84	0.76	0.62
	DT score	0.81	0.76	0.75	0.75	0.76	0.54
	DT case	0.78	0.79	0.77	0.79	0.69	0.46
	RHS case	0.76	0.81	0.77	0.82	0.73	0.58
	13-Item score	0.91	0.90	0.88	0.89	0.89	0.67
	13-Item case	0.77	0.84	0.79	0.84	0.76	0.62
Burma	Total score	0.49	0.13	0.56	0.14	0.72	0.54
	Total case	0.31	-0.05	0.24	-0.05	0.65	0.52
	DT score	0.52	0.17	0.58	0.18	0.62	0.49
	DT case	0.34	-0.02	0.23	-0.03	0.44	0.56
	RHS case	0.31	-0.05	0.24	-0.05	0.65	0.52
	13-Item score	0.55	0.17	0.58	0.18	0.73	0.54
	13-Item case	0.49	0.36	0.50	0.39	0.53	0.40
Bhutan	Total score	0.83	0.73	0.84	0.70	0.92	0.77
	Total case	0.74	0.80	0.76	0.74	0.83	0.93
	DT score	0.46	0.51	0.52	0.46	0.45	0.43
	DT case	0.33	0.37	0.41	0.37	0.35	0.34
	RHS case	0.58	0.59	0.66	0.62	0.66	0.68
	13-Item score	0.84	0.74	0.84	0.70	0.91	0.77
	13-Item case	0.69	0.71	0.73	0.66	0.79	0.85

See definitions for RHS-15 components and DPs in footnotes to Tables 1 and 2.

Predictive Validity: Sensitivity and Specificity of Various Cutoff Scores for Total Score and 13-Item Score on DPs

To facilitate possible adaptation of the RHS-15 for different clinical sites where resources may vary, Table 5 shows the sensitivity and specificity of various cutoff scores of the total score and the 13-item score for each DP. At each cutoff score (x), the 13-item score equivalent (x - 1) shows equal or slightly less sensitivity, with the exception of the 16-15 cutoff score pair for depression. Specificity is variable but similar for each cutoff score equivalent for each DP.

DISCUSSION

These data corroborate earlier findings that the RHS-15 is a useful and valid screening instrument for detecting common emotional disorders in a combined sample of 3 refugee groups. This study addresses previous limitations of no data about the efficacy of the RHS-15 in a sample being assessed in a real-world public health setting (Hollifield et al., 2013) and further assesses if and how the RHS-15 may be more efficient. While the current RHS-15 has acceptable metric properties, adaptation to a 13-item instrument may improve efficiency and patient and provider acceptance without compromising sensitivity and specificity. Possible study limitations include a relatively low case rate for the Burmese, the use of DPs instead of clinical diagnoses, and limits on reliability of data, particularly of the Burmese population. Diagnostic proxies used have been validated in refugees. It is not known how the RHS-15 would perform using diagnostic interviews as the measurement standard. It is also not clear that such diagnostic interviews would be any more of a relevant standard than the proxies used. A larger sample size may have enhanced power to assess validity by country of origin and to more reliably assess factor structure of the RHS-15 components by country. However, while there are rules of thumb that suggest the need for 5, 10, or 20 individuals for each variable to conduct principal component analysis, those rules are not well studied. It is probably more accurate to say that “strong data” in factor analysis mean consistently high communalities (individual 0.6 and mean 0.7) without cross loadings and several variables loading strongly on each factor. If these conditions are true, and if the number of expected factors is relatively small, and model error is low (which often goes hand-in-hand with high communalities), researchers and reviewers should not be overly concerned about small sample sizes (Preacher and MacCallum, 2002). Total sample and Iraqi and Nepali data conformed to these conditions, but Burmese data did not. We are able to thus conclude that the RHS-15 and the 13-item components appear to be 1 general measure of distress, and more study is needed in other populations to evaluate construct validity.

While the RHS-15 seems valid for detecting 3 common mental disorders, it should be viewed as a general screening instrument to be used in the public health setting. Just as diabetes is not diagnosed with 1 blood glucose screening, psychiatric disorders are not intended to be diagnosed with the RHS-15. Referral for culturally informed individualized assessment is required. This process is perhaps even more crucial in refugee populations. Western diagnostic nosology has significant limitations in general and even more so with refugee populations because of variation in causality, sociocultural context, and symptom manifestation (deGirolo and McFarlane, 1996; Hollifield et al., 2003; Punakami, 1989). The literature provides many examples of how sociocultural variables are associated with misdiagnosis and improper treatment, especially in minority and vulnerable populations (Good, 1992; Neighbors et al., 1989). There is an extensive literature describing how culture influences patients’ experience of psychopathology, the clinical manifestations and course of psychiatric illness, diagnosis, treatment, and treatment outcomes (Fabrega, 1987; Kleinman, 1988; Mezzich, 1984; Rogler, 1989). Furthermore, for those having experienced war, oppression, migration, and resettlement, it is usual to suffer a broad range of symptoms (Hollifield et al., 2009), qualifying

TABLE 4. Predictive Validity of Total Case, RHS Case, and 13-Item Case to Diagnostic Proxies

		ANX		DEP		PTSD	
		Pos	Neg	Pos	Neg	Pos	Neg
Total case	Pos	49	19	52	16	58	10
	Neg	1	104	4	100	9	96
	Stat	$\chi^2 = 101.5; p < 0.0001$		$\chi^2 = 98.8; p < 0.0001$		$\chi^2 = 102.4; p < 0.0001$	
	PPV	0.72		0.76		0.85	
	NPV	0.99		0.96		0.91	
	Sens/spec	0.98/0.85		0.93/0.86		0.87/0.91	
RHS case	Pos	50	28	55	23	59	19
	Neg	1	95	2	93	9	87
	Stat	$\chi^2 = 82.6; p < 0.0001$		$\chi^2 = 90.7; p < 0.0001$		$\chi^2 = 79.4; p < 0.0001$	
	PPV	0.64		0.71		0.76	
	NPV	0.99		0.98		0.91	
	Sens/spec	0.98/0.77		0.96/0.80		0.87/0.82	
13-Item case	Pos	48	17	51	14	55	10
	Neg	2	106	5	102	12	96
	Stat	$\chi^2 = 102.4; p < 0.0001$		$\chi^2 = 100.3; p < 0.0001$		$\chi^2 = 92.4; p < 0.0001$	
	PPV	0.74		0.78		0.85	
	NPV	0.98		0.95		0.89	
	Sens/spec	0.96/0.86		0.91/0.88		0.82/0.91	

See definitions for RHS-15 components and DPs in footnotes to Tables 1 and 2.

for multiple diagnoses. A recent study of all refugee patients at the Competence Center for Transcultural Psychiatry in Copenhagen showed that almost all patients had depression, pain, and untreated somatic complaints in addition to PTSD. Furthermore, 36% to 58% were in treatment for physical problems, 9% to 16% had psychotic symptoms, 27% had enduring personality changes, and 46% reported traumatic brain injury (Buhmann, 2014). Thus, a screening instrument ought to

be sensitive for identifying those with a broad range of symptoms and potential diagnoses.

Current users of the RHS-15 may continue to implement it with confidence in its usefulness. If efficiency and time are key considerations, removal of either the coping item (item 14) and/or the DT (item 15) appears acceptable from a metric perspective. The coping item is the least correlated with the other items and the worst fit in factor

TABLE 5. Sensitivity and Specificity at Various Cutoff Scores for Total Score and 13-Item Score

Sensitivity and Specificity to Diagnostic Proxies at Different Cutoff Scores, Total n = 179

Proxy Diagnosis		RHS-15 Total Score (Items 1–14)									
		9	10	11	12	13	14	15	16	17	18
PTSD	Sensitivity	0.925	0.910	0.866	0.866	0.836	0.836	0.776	0.731	0.731	0.702
	Specificity	0.793	0.859	0.868	0.906	0.906	0.925	0.934	0.953	0.972	0.972
ANX	Sensitivity	1.000	1.000	1.000	0.980	0.960	0.960	0.880	0.880	0.860	0.820
	Specificity	0.724	0.789	0.821	0.846	0.854	0.870	0.878	0.919	0.927	0.927
DEP	Sensitivity	0.964	0.964	0.964	0.929	0.911	0.911	0.857	0.839	0.821	0.804
	Specificity	0.793	0.859	0.868	0.906	0.906	0.925	0.934	0.953	0.972	0.972

Sensitivity and Specificity to Diagnostic Proxies at Different Cutoff Scores, Total n = 179

Proxy Diagnosis		RHS-15 13-Item Score (Items 1–13)									
		9	10	11	12	13	14	15	16	17	18
PTSD	Sensitivity	0.881	0.851	0.821	0.821	0.761	0.746	0.746	0.702	0.702	0.687
	Specificity	0.859	0.887	0.906	0.915	0.915	0.943	0.953	0.972	0.972	0.972
ANX	Sensitivity	1.000	1.000	0.960	0.940	0.880	0.880	0.880	0.840	0.840	0.820
	Specificity	0.805	0.846	0.862	0.862	0.870	0.902	0.911	0.935	0.935	0.935
DEP	Sensitivity	0.964	0.946	0.911	0.893	0.857	0.857	0.857	0.804	0.804	0.804
	Specificity	0.859	0.887	0.906	0.915	0.915	0.943	0.953	0.972	0.972	0.972

See definitions for RHS-15 components and DPs in footnotes to Tables 1 and 2.

analyses. The inclusion of the DT reduces specificity (ie, more cases are apt to be false positives) compared with total score or 13-item score, which is slightly less sensitive and similar in specificity to DPs than total score. The DT, however, may be the most useful in the Burmese. Given the study findings, P2W recommends that the 13-item score (≥ 11) can be used as the primary case identifier in settings where the RHS-15 is used solely as a screening instrument. In clinical settings, the coping item and DT may be used to further contextualize the case. For example, if the 13-item score is less than the recommended cutoff score of 11 and the DT is 5 or greater, utilize a follow-up question, such as “You marked on the thermometer that you feel quite (very) distressed; what is the source of your distress?” The clinician might determine if the source of distress is symptoms, or alternatively if the source is family, social, or resettlement concerns, which may help with treatment planning. Likewise, the coping question might be used when a person scores close to the cutoff score by probing for their need for services based on the coping response. Also, a clinical site may choose a different cutoff score, higher if the preference is for more specificity because there are fewer services available and lower for more sensitivity if service limitation is not an obstacle. If this is done, the procedure should be defined and well documented with a standard protocol. The RHS will ideally be used to optimize sensitivity and to plan for a second interview to determine the acuity of distress (ie, immediate risk to self or others) and if treatment for the individual is desired and warranted. This second-level interview should be conducted by clinicians who are trained in culturally informed diagnosis and treatment planning. Importantly, this interview would include evaluation of suicidality, psychosis, substance use, and potential violence because the RHS-15 was intentionally developed to be a screening instrument and not a substitute for an important clinical evaluation.

The recommendation to optimize for sensitivity is consistent with principles of screening for potentially serious medical conditions. Emotional distress that is predictive of psychiatric disorder is such a condition. Rates of suicidal ideation and attempts in those with anxiety and mood disorders are higher than those in the general population, and emotional distress is a primary variable associated with suicide behaviors, which varies by ethnicity (Kessler et al., 2005; Oquendo et al., 2001). Treating depression reduces suicide risk (Rihmer, 2001). Anxiety disorders are associated with higher disability, use of health services, substance use, and financial dependency (Hollifield et al., 1997). Refugees are at risk of decreased life expectancy and worse overall health associated with trauma and distress (Hollifield et al., 2002; Wong et al., 2011). And, not detailed here but rather the topic of an upcoming report is that the prevalence of distress in refugees at the civil surgeon examination is higher than at the incoming domestic medical examination 8 to 12 months earlier. The high rates of ongoing distress and psychiatric disorder coupled with the associated morbidity and mortality make screening for distress in newly arrived refugees and at later time points a critical medical and ethical issue.

Implementation of screening is critical to its efficacy and effectiveness. Public health and clinical staff must be motivated to conduct screening properly and sensitively. Developing rapport and diminishing stigma, having a method of referral, and fidelity of administration are important to reduce variability in case finding and improve overall effectiveness of screening. To enhance rapport and diminish stigma, P2W developed the RHS-15 to be administered with routine medical screening using neutral medical language (ie, not referring to “mental health”), utilizing scripts to discuss a positive case and referral for care. These methods during the initial RHS-15 development resulted in 74% referral acceptance and 62% engagement in care for positive cases. Partners at other sites have found rates of referral acceptance and engagement in care to be less than this. At the time of this writing, the RHS-15 has been translated to 12 languages, including Amharic, Arabic, Burmese, Farsi, French, Karen, Nepali, Russian, Somali, Spanish (Cuban version), Swahili, and Tigrinya. Translation was standardized in

order to include culturally specific terms for important concepts, approved by bilingual committees from each ethnic group. Using these translated versions, whether self-administered or translator administered is important to maintain fidelity. Finally, while all newly arrived refugees should for medical and ethical reasons be screened for distress, we caution against proscriptive rules and regulations for screening and referring that are bureaucratic and not clinical. For example, mandating a “mental health” referral for someone who screens positive could potentially do more harm than good (Pottie et al., 2011). It is critical to screen incoming refugees for distress and offer help in the context of expert and culturally informed practice. Further knowledge about best practices for implementation will be guided by ongoing research about the RHS-15 in different languages and health care settings.

CONCLUSIONS

The RHS-15 is a practical and useful instrument for screening refugees for emotional distress in a public health setting. It is currently translated into 12 languages and is being used at many sites across the world. More research about reliability in various ethnic groups is needed. The original 15-item instrument has similar metric properties to a shorter, 13-item tool. While full assessment, diagnosis, and treatment planning are complex, particularly for various populations, the RHS-15 may be used to conduct critical screening for emotional distress.

ACKNOWLEDGMENTS

The authors appreciate the multiple translators, community refugee focus group participants, data assistance of Ryan Doria, and manuscript preparation assistance of Andrea Gory. They also appreciate their collaborators across the world who are utilizing and further evaluating the RHS-15. They thank the Refugee Health Technical Assistance Center, the Center for Victims of Torture, and the North American Refugee Health Conference for featuring this work in virtual or on-site conferences.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

- Brislin RW (1970) Back-translation for cross-cultural research. *J Cross Cult Psychol*. 1:187–216.
- Buhmann CB (2014) Traumatized refugees: Morbidity, treatment and predictors of outcome. *Dan Med J*. 61:B4871.
- Butcher JN (1991) Psychological evaluation. In Westermeyer J, Williams C, Nguyen A (Eds), *Mental health services for refugees* (pp 111–122). Washington, DC: Government Printing Office.
- CDC (Accessed July 27, 2015) Guidelines for mental health screening during the domestic medical examination for newly arrived refugees. <http://www.cdc.gov/immigrantrefugeehealth/guidelines/domestic/mental-health-screening-guidelines.html>.
- deGirolo GD, McFarlane AC (1996) The epidemiology of PTSD: A comprehensive review of the international literature. In Marsella A, Friedman M, Gerrity E (Eds), *Ethnocultural aspects of posttraumatic stress disorder* (pp 93–119). Washington, DC: American Psychological Association.
- de Jong JT, Komproe IH, van Ommeren M, El Masri M, Araya M, Khaled N, van de Put W, Somasundaram D (2001) Lifetime events and posttraumatic stress disorder in 4 postconflict settings. *JAMA*. 286:555–562.
- Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L (1974) The Hopkins Symptom Checklist (HSCL): A measure of primary symptom dimensions. In Basel P (Ed), *Modern problems in pharmacopsychiatry*. Basel, Switzerland: Karger.
- Fabrega HJ (1987) Psychiatric diagnosis. A cultural perspective. *J Nerv Ment Dis*. 175: 383–394.
- Fazel M, Wheeler J, Danesh J (2005) Prevalence of serious mental disorder in 7000 refugees resettled in western countries: A systematic review. *Lancet*. 365:1309–1314.

- Fenta H, Hyman I, Noh S (2004) Determinants of depression among Ethiopian immigrants and refugees in Toronto. *J Nerv Ment Dis.* 192:363–372.
- Foa EB, Riggs DS, Dancu CV, Rothbaum BO (1993) Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *J Trauma Stress.* 6:459–473.
- Good BJ (1992) Culture, diagnosis and comorbidity. *Cult Med Psychiatry.* 16:427–446.
- Hollifield M, Geppert C, Johnson Y, Fryer C (2003) A Vietnamese man with selective mutism: The relevance of multiple interacting 'cultures' in clinical psychiatry. *Transcult Psychiatry.* 40:329–341.
- Hollifield M, Katon W, Skipper B, Chapman T, Ballenger JC, Fyer A (1997) Panic disorder and quality of life: Variables predictive of functional impairment. *Am J Psychiatry.* 154:766–772.
- Hollifield M, Verbillis-Kolp S, Farmer B, Toolson EC, Woldehaimanot T, Yamazaki J, Holland A, St Clair J, Soohoo J (2013) The Refugee Health Screener-15 (RHS-15): Development and validation of an instrument for anxiety, depression, and PTSD in refugees. *Gen Hosp Psychiatry.* 35:202–209.
- Hollifield M, Warner T, Jenkins J, Sinclair-Lian N, Krakow B, Eckert V, Karadaghi P, Westermeyer J (2006) Assessing war trauma in refugees: Properties of the Comprehensive Trauma Inventory-104 (CTI-104). *J Trauma Stress.* 19:527–540.
- Hollifield M, Warner TD, Krakow B, Jenkins J, Westermeyer J (2009) The range of symptoms in refugees of war: The New Mexico Refugee Symptom Checklist-121. *J Nerv Ment Dis.* 197:117–125.
- Hollifield M, Warner TD, Lian N, Krakow B, Jenkins JH, Kesler J, Stevenson J, Westermeyer J (2002) Measuring trauma and health status in refugees: A critical review. *JAMA.* 288:611–621.
- Jaranson JM, Butcher J, Halcon L, Johnson DR, Robertson C, Savik K, Spring M, Westermeyer J (2004) Somali and Oromo refugees: Correlates of torture and trauma history. *Am J Public Health.* 94:591–598.
- Kessler RC, Berglund P, Borges G, Nock M, Wang PS (2005) Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990–1992 to 2001–2003. *JAMA.* 293:2487–2495.
- Kinzie JD, Manson SM (1987) The use of self-rating scales in cross-cultural psychiatry. *Hosp Community Psychiatry.* 38:190–196.
- Kleinman A (1988) *Rethinking psychiatry: From cultural category to personal experience.* New York: Free Press.
- Laban CJ, Gemaat HB, Komproe IH, van der Tweel I, de Jong JT (2004) Postmigration living problems and common psychiatric disorders in Iraqi asylum seekers in the Netherlands. *J Nerv Ment Dis.* 193:825–832.
- Marshall GN, Schell TL, Elliott MN, Berthold SM, Chun C (2005) Mental health of Cambodian refugees two decades after resettlement in the United States. *JAMA.* 294:571–579.
- Mezzich JE (1984) *Culture and psychopathology.* New York: Columbia University Press.
- Mollica R, McInnes K, Pham T, Smith Fawzi M, Murphy E, Lin L (1998) The dose-effect relationships between torture and psychiatric symptoms in Vietnamese ex-political detainees and a comparison group. *J Nerv Ment Dis.* 186:543–553.
- Mollica RF, Wyshak G, de Marnette D, Khuon F, Lavelle J (1987) Indochinese versions of the Hopkins Symptom Checklist-25: A screening instrument for the psychiatric care of refugees. *Am J Psychiatry.* 144:497–500.
- Momartin S, Steel Z, Coello M, Aroche J, Silove D, Brooks R (2006) A comparison of the mental health of refugees with temporary versus permanent protection visas. *Med J Aust.* 185:357–361.
- Neighbors HW, Jackson JS, Campbell L, Williams D (1989) The influence of racial factors on psychiatric diagnosis: A review and suggestions for research. *Community Ment Health J.* 25:301–311.
- Oquendo MA, Ellis SP, Greenwald S, Malone KM, Weissman MM, Mann JJ (2001) Ethnic and sex differences in suicide rates relative to major depression in the United States. *Am J Psychiatry.* 158:1652–1658.
- Porter M, Haslam N (2005) Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: A meta-analysis. *JAMA.* 294:602–612.
- Pottie K, Greenaway C, Feightner J, Welch V, Swinkels H, Rashid M, Narasiah L, Kirmayer LJ, Ueffing E, MacDonald NE, Hassan G, McNally M, Khan K, Buhrmann R, Dunn S, Dominic A, McCarthy AE, Gagnon AJ, Rousseau C, Tugwell P (2011) Evidence-based clinical guidelines for immigrants and refugees. *CMAJ.* 183:E824–E925.
- Preacher KJ, MacCallum RC (2002) Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. *Behav Genet.* 32:153–161.
- Punakami RI (1989) Political violence and mental health. *Int J Ment Health.* 17:3–15.
- Rihmer Z (2001) Can better recognition and treatment of depression reduce suicide rates? A brief review. *Eur Psychiatry.* 16:406–409.
- Rogler LH (1989) The meaning of culturally sensitive research in mental health. *Am J Psychiatry.* 146:296–303.
- Savin D, Seymour DJ, Littleford LN, Bettridge J, Giese A (2005) Findings from mental health screening of newly arrived refugees in Colorado. *Public Health Rep.* 120:224–229.
- Services, Bridging Refugee Youth and Children's Services (September 2011). <http://www.brycs.org/aboutRefugees/refugee101.cfm>.
- Shannon P, Im H, Becher E, Simmilink J, Weiling E, O'Fallon A (2012) Screening for war trauma, torture, and mental health symptoms among newly-arrived refugees: A national survey of U.S. refugee health coordinators. *J Immigr Refugee Stud.* 10:380–394.
- United Nations High Commissioner for Refugees. <http://www.unhcr.org.uk/about-us/key-facts-and-figures.html>. Accessed August 23, 2015.
- US Committee for Refugees and Immigrants World Refugee Survey (2009). <http://www.refugees.org/resources/refugee-warehousing/archived-world-refugee-surveys/2009-world-refugee-survey.html>. Accessed December 30, 2014.
- Weine SM, Kulenovic AD, Pavkovic I, Gibbons R (1998) Testimony psychotherapy in Bosnian refugees: A pilot study. *Am J Psychiatry.* 155:1720–1726.
- Westermeyer J, Janca A (1997) Language, culture and psychopathology: Conceptual and methodological issues. *Transcult Psychiatry.* 34:291–311.
- Winokur A, Winokur DF, Rickels K, Cox D (1984) Symptoms of emotional distress in a family planning service: Stability over a four-week period. *Br J Psychiatry.* 144:395–399.
- Wong EC, Marshall GN, Schell TL, Elliott MN, Babey SH, Hambarsoomians K (2011) The unusually poor physical health status of Cambodian refugees two decades after resettlement. *J Immigr Minor Health.* 13:876–882.