

# The SF-36 Offers a Strong Measure of Mental Health Symptoms in Survivors of Acute Respiratory Failure

## A Tri-National Analysis

Elizabeth R. Pfoh<sup>1</sup>, Kitty S. Chan<sup>2</sup>, Victor D. Dinglas<sup>3,4</sup>, Brian H. Cuthbertson<sup>5,6</sup>, Doug Elliott<sup>7</sup>, Richard Porter<sup>8</sup>, O. Joseph Bienvenu<sup>9</sup>, Ramona O. Hopkins<sup>10,11,12</sup>, and Dale M. Needham<sup>3,4,13</sup>

<sup>1</sup>Division of General Internal Medicine, <sup>3</sup>Outcomes after Critical Illness and Surgery Group, <sup>4</sup>Division of Pulmonary and Critical Care Medicine, <sup>9</sup>Department of Psychiatry and Behavioral Sciences, and <sup>13</sup>Department of Physical Medicine and Rehabilitation, Johns Hopkins School of Medicine, Baltimore, Maryland; <sup>2</sup>Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; <sup>5</sup>Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; <sup>6</sup>Department of Anesthesia, University of Toronto, Toronto, Ontario, Canada; <sup>7</sup>Faculty of Health, University of Technology, Sydney, Australia; <sup>8</sup>Adult Intensive Care Unit, Glenfield Hospital, Leicester, United Kingdom; <sup>10</sup>Division of Pulmonary and Critical Care Medicine, Intermountain Medical Center, Murray, Utah; <sup>11</sup>Psychology Department and Neuroscience Center, Brigham Young University, Provo, Utah; and <sup>12</sup>Center for Humanizing Critical Care, Intermountain Health Care, Murray, Utah

ORCID ID: 0000-0003-2639-4186 (K.S.C.).

### Abstract

**Rationale:** Survivors of acute respiratory failure commonly experience long-term psychological sequelae and impaired quality of life. For researchers interested in general mental health, using multiple condition-specific instruments may be unnecessary and inefficient when using the Medical Outcomes Study Short Form (SF)-36, a recommended outcome measure, may suffice. However, relationships between the SF-36 scores and commonly used measures of psychological symptoms in acute survivors of respiratory failure are unknown.

**Objectives:** Our objective is to examine the relationship of the SF-36 mental health domain (MH) and mental health component summary (MCS) scores with symptoms of depression, anxiety, and post-traumatic stress disorder (PTSD) evaluated using validated psychological instruments.

**Methods:** We conducted a cross-sectional analysis of 1,229 participants at 6- and 12-month follow-up assessment using data from five studies from the United States, the United Kingdom, and Australia.

**Measurements and Main Results:** Symptoms were assessed using the Hospital Anxiety and Depression Scale (HADS), Depression Anxiety Stress Scales, the Davidson Trauma Scale, Impact of Event

Scale (IES), and IES-Revised (IES-R). At 6-month assessment there were moderate to strong correlations of the SF-36 MH scores with HADS depression and anxiety symptoms ( $r = -0.74$  and  $-0.79$ ) and with IES-R PTSD symptoms ( $r = -0.60$ ) in the pooled analyses. Using the normalized population mean of 50 on the SF-36 MH domain score as a cut-off, positive predictive values were 16 and 55% for substantial depression; 20 and 68% for substantial anxiety (Depression Anxiety Stress Scales and HADS, respectively); and 40, 44, and 67% for substantial PTSD symptoms (IES-R, IES, and Davidson Trauma Scale, respectively). Negative predictive values were high. The area under the receiver operating characteristics curve of the SF-36 MH score was high for depression, anxiety, and PTSD symptoms (0.88, 0.91, and 0.84, respectively). All results were consistent for the MCS, across the individual studies, and for the 12-month assessment.

**Conclusions:** For researchers interested in general mental health status, the SF-36 MH or MCS offers a strong measure of psychological symptoms prevalent among survivors of acute respiratory failure. For researchers interested in specific conditions, validated psychological instruments should be considered.

**Keywords:** acute respiratory failure; Medical Outcomes Study Short Form 36 survey; depression; anxiety; post-traumatic stress disorder

(Received in original form October 20, 2015; accepted in final form February 16, 2016)

Supported by National Institutes of Health grants R24HL111895, R01HL091760, R01HL091760-02S1, R01HL096504, K23AG034257, HSN268200536170C, HHSN268200536171C, HHSN268200536173C, HHSN268200536174C, HSN268200536175C, and HHSN268200536179C; the Johns Hopkins Institute for Clinical and Translational Research grant UL1 TR 000424-06; Health Resources and Services Administration grant T32HP10025B0; and the Australian National Health and Medical Research Council.

The sponsor had no role in the data acquisition, analysis, or preparation of the manuscript.

**Author Contributions:** E.R.P. takes primary responsibility for the content of this manuscript including the data and analysis. All authors contributed to the conception and/or design of this study. D.M.N., V.D.D., B.H.C., D.E., R.P., O.J.B., and R.O.H. contributed to the acquisition of data. E.R.P. and K.S.C. contributed to the analysis of data, and all authors contributed to the interpretation of data. E.R.P. and D.M.N. drafted the manuscript, and all authors critically revised it for important intellectual content and approved the final version to be submitted.

Correspondence and requests for reprints should be addressed to Elizabeth R. Pfoh, Ph.D., M.P.H., Research Fellow, Division of General Internal Medicine, Johns Hopkins University, 2024 East Monument Street, Suite 2-516F, Baltimore, MD 21208. E-mail: epfoh1@jhu.edu

This article has an online supplement, which is accessible from this issue's table of contents at [www.atsjournals.org](http://www.atsjournals.org)

Ann Am Thorac Soc Vol 13, No 8, pp 1343–1350, Aug 2016

Copyright © 2016 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201510-705OC

Internet address: [www.atsjournals.org](http://www.atsjournals.org)

Survivors of acute respiratory failure requiring mechanical ventilation in an intensive care unit (ICU) commonly experience long-term psychological sequelae and impaired quality of life (1–3). In studies evaluating survivors of acute respiratory failure, quality of life status after hospital discharge is a recommended outcome measure (4), and the Medical Outcomes Study Short Form 36 (SF-36) instrument is commonly used and recommended (5–7). The SF-36 is a self-report generic measure of health-related quality of life (8) that has been validated in survivors of acute respiratory failure (9, 10). The SF-36 has a 5-item mental health domain (MH) score and a 35-item overall mental health component summary score (MCS) that evaluate general mental health status (11).

Researchers interested in specific psychological sequela of survivors of acute respiratory failure, such as depression, anxiety and post-traumatic stress disorder (PTSD) symptoms, must use multiple assessment instruments. However, not all researchers are interested in specific psychological conditions, and some may only be interested in general mental health status. For these researchers, using multiple condition-specific instruments may be an unnecessary burden on research participants and an inefficient use of research resources, and the SF-36's MH or MCS scores may suffice. However, relationships between the SF-36 scores and commonly used measures of psychological symptoms in patients with acute respiratory failure are unknown. Such knowledge may aid in designing a minimum set of core outcome measures that should be included in all clinical trials (i.e., a core outcome set) (12) evaluating postdischarge patient outcomes to improve the comparability of outcomes across trials (13). Core outcome sets, with limited redundancy in patient outcome measures, would reduce cost and patient burden and improve feasibility of conducting outcome assessments after hospital discharge.

To provide data to help inform the development of a minimum set of outcome measures to be used in all long-term outcomes research in this field (i.e., a core outcome set) (12, 14), we undertook this study to examine the relationship of the SF-36 MH and MCS scores with symptoms of depression, anxiety, and PTSD, evaluated

using common psychological self-report instruments, in survivors of acute respiratory failure. Secondly, because we are evaluating use of the SF-36 as a measure of general mental health, we also confirmed the SF-36 factor structure in survivors of acute respiratory failure, because there are limited data confirming that its main constructs (i.e., physical and mental health) are similar in survivors of acute respiratory failure versus the general population.

## Methods

### Study Design and Data Sources

We conducted a cross-sectional analysis of data collected at 6- and/or 12-month follow-up assessments for survivors of acute respiratory failure enrolled in five studies conducted in the United States (ARDSNet Long Term Outcomes Study [ALTOS] and Improving Care of Acute Lung Injury Patients [ICAP] studies), United Kingdom (Conventional ventilatory support vs. extracorporeal membrane oxygenation for severe adult respiratory failure [CESAR] and a pragmatic randomized, controlled trial of intensive care follow-up programs in improving longer-term outcomes from critical illness [PRaCTICAL] studies), and Australia (Elliott and colleagues study [15]). All studies received ethics approval and obtained informed consent. Each study is briefly described herein.

ALTOS included patients in the United States who were enrolled in three of the ARDS Network's randomized trials, with 6- and 12-month assessments occurring between 2008 and 2012 (16–19). The three randomized trials included 41 hospitals. The ICAP study was a prospective cohort study evaluating acute respiratory distress syndrome survivors from four hospitals in Baltimore, Maryland with 6- and 12-month assessments occurring between 2005 and 2008 (20). The CESAR study was a randomized control trial of conventional ventilation versus extracorporeal membrane oxygenation for severe acute respiratory failure that enrolled patients from 103 hospitals in England with a 6-month assessment occurring between 2002 and 2007 (21). The PRaCTICAL study was a pragmatic randomized trial that enrolled patients from three hospitals in the United Kingdom with 6- and 12-month

assessments occurring between 2007 and 2008 (22, 23). Finally, the Elliott and colleagues study was a randomized trial of 12 hospitals in Australia with a 6-month assessment occurring between 2005 and 2009 (15). To be eligible for the current analysis, a patient must have completed the SF-36 at the 6- and/or 12-month assessment.

### Measures

SF-36 version 2 (24, 25) is the outcome measure being evaluated in this study. Psychological symptoms were assessed with commonly used self-report instruments for depression, anxiety, and PTSD symptoms. Specifically, the Hospital Anxiety and Depression Scale (HADS) (16, 26) and Depression Anxiety Stress Scales (DASS) (27, 28) assessed depression and anxiety. The Davidson Trauma Scale (DTS) (29), Impact of Event Scale (IES) (30), and IES-Revised (IES-R) (31) assessed PTSD. For depression and anxiety symptoms, the HADS was used by the ICAP, ALTOS, CESAR, and PRaCTICAL studies, and the DASS was used by the Elliott and colleagues study. For PTSD symptoms, the IES-R was used by ICAP and ALTOS, the DTS by PRaCTICAL, and the IES by the Elliott and colleagues study (15).

For this analysis, psychological symptoms were evaluated as binary outcomes: presence or absence of substantial symptoms using traditional instrument-specific threshold (i.e.,  $\geq 8$  for HADS depression and anxiety subscales [16, 26];  $\geq 21$  for DASS depression [27, 28];  $\geq 15$  for DASS anxiety [27, 28];  $\geq 27$  for DTS [23, 29];  $\geq 20$  for IES [30, 32]; and  $\geq 1.6$  for IES-R [33]).

### Statistical Analysis

In addition to descriptive statistics, the associations between continuous scores from each psychological instrument and the SF-36 MH and MCS scores were examined using Spearman correlation coefficients. We defined the strength of the correlation as "very strong" if rho was greater than or equal to |0.80|, "strong" if rho was greater than or equal to |0.60| and less than or equal to |0.79|, "moderate" if rho was greater than or equal to |0.40| and less than or equal to |0.59|, and "weak" if rho was less than |0.39| (34). The area under receiver operating characteristics (AUROC) curves was used to evaluate the

discrimination of the SF-36 score for survivors who did versus did not have substantial symptoms of each of the psychological outcomes (depression, anxiety, and PTSD). The AUROC curves were calculated for both the MCS and MH scores. However, because the results were similar and not all of the studies collected the MCS results, we only reported the MH scores. Sensitivity, specificity, and positive and negative predictive values for SF-36 normalized scores (mean, 50; SD, 10) of 40, 45, 50, 55, and 60 were calculated to identify SF-36 thresholds that may indicate the presence of substantial psychological symptoms. When possible, studies using the same psychological instrument were pooled

using patient-level data from each study. Finally, we conducted a sensitivity analysis of the pooled correlation analysis removing the CESAR dataset from the pooled data.

Exploratory factor analysis of the SF-36 at 6- and 12-month assessments was performed using data from studies that had the SF-36 MCS available. Only one factor had an eigenvalue greater than 1 at the 6- and 12-month time points, suggesting that one factor should be extracted. Scree plots of the eigenvalues supported extracting one factor. Parallel analysis, which fits multiple “slices” of data with different weights to identify the number of factors (35), indicated that three factors should be extracted. On the basis of conventional

criteria for determining the number of factors—eigenvalues, scree plots, parallel analysis, and theory—we specified that two factors (physical and mental health) be extracted. We used orthogonal and oblique factor rotations to compare results from our population to a nationally representative U.S. sample (36), which was similar to prior research using samples from the United Kingdom (37, 38) and Australia (39). All analyses were performed using Stata 13.0.

Results

In total, 1,229 survivors of acute respiratory failure were included across the five

Table 1. Participant characteristics at baseline, 6 months, and 12 months, by study

	ICAP (20), U.S. (N = 164)	ALTOS (16), U.S. (N = 613)	Elliott <i>et al.</i> (15), Australia (N = 159)	CESAR (21), UK (N = 74)	PRaCTICAL (22), UK (N = 219)
Patient population	ARDS	ARDS	ARF	Severe ARF	ARF
Age, yr	48 (14)	49 (14)	57 (16)	38 (13)	58 (15)
Men, n (%)	91 (56)	309 (49)	98 (62)	43 (58)	122 (56)
APACHE II	24 (8)	26 (8)*	20 (10)	19 (6)	19 (7)
Ventilation days	14 (14)	11 (10)	6 (6)	—	7 (8)
ICU length of stay	19 (17)	14 (11)	9 (8)	—	7 (8)
Hospital length of stay	31 (23)	22 (16)	24 (19)	63 (49)	—
6-mo values					
SF-36 MH	47 (13)	44 (14)	48 (11)	46 (12)	46 (13)
SF-36 MCS	48 (14)	45 (15)	47 (12)	—	45 (13)
Depression, <sup>†</sup> n (%)	41 (25) <sup>‡</sup>	222 (36) <sup>‡</sup>	15 (9) <sup>§</sup>	16 (22) <sup>‡</sup>	60 (27) <sup>‡</sup>
Anxiety, <sup>†</sup> n (%)	52 (33) <sup>‡</sup>	260 (42) <sup>‡</sup>	18 (11) <sup>§</sup>	25 (34) <sup>‡</sup>	86 (39) <sup>‡</sup>
PTSD, <sup>†</sup> n (%)	31 (19) <sup>  </sup>	148 (24) <sup>  </sup>	46 (29) <sup>  </sup>	—	92 (42) <sup>**</sup>
Depression and anxiety, n (%)	29 (18) <sup>‡</sup>	166 (27) <sup>‡</sup>	11 (7) <sup>§</sup>	14 (19) <sup>‡</sup>	48 (22) <sup>‡</sup>
Depression and PTSD, n (%)	17 (11) <sup>‡,  </sup>	117 (19) <sup>‡,  </sup>	11 (7) <sup>§,  </sup>	—	46 (21) <sup>‡,***</sup>
Anxiety and PTSD, n (%)	20 (13)	128 (21)	13 (8)	—	63 (29)
12-mo values					
SF-36 MH	47 (13)	45 (14)	—	—	48 (12)
SF-36 MCS	47 (14)	45 (15)	—	—	47 (13)
Depression, <sup>†</sup> n (%)	34 (24) <sup>‡</sup>	204 (36) <sup>‡</sup>	—	—	53 (27) <sup>‡</sup>
Anxiety, <sup>†</sup> n (%)	50 (35) <sup>‡</sup>	241 (42) <sup>‡</sup>	—	—	68 (35) <sup>‡</sup>
PTSD, <sup>†</sup> n (%)	31 (22) <sup>  </sup>	132 (23) <sup>  </sup>	—	—	77 (40) <sup>**</sup>
Depression and anxiety, n (%)	24 (17) <sup>‡</sup>	158 (28) <sup>‡</sup>	—	—	40 (21) <sup>‡</sup>
Depression and PTSD, n (%)	18 (13) <sup>‡,  </sup>	93 (16) <sup>‡,  </sup>	—	—	45 (23) <sup>‡,***</sup>
Anxiety and PTSD, n (%)	22 (16) <sup>‡,  </sup>	117 (20) <sup>‡,  </sup>	—	—	56 (29) <sup>‡,***</sup>

Definition of abbreviations: ALTOS = ARDSNet Long Term Outcomes study; APACHE = Acute Physiology and Chronic Health Evaluation; ARDS = acute respiratory distress syndrome; ARF = acute respiratory failure; CESAR = conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure; DASS = Depression Anxiety Stress Scales; DTS = Davidson Trauma Scale; HADS = Hospital Anxiety and Depression Scale; ICAP = Improving Care of Acute Lung Injury Patients study; ICU = intensive care unit; IES = Impact of Event Scale; IES-R = Impact of Event Scale-Revised; MCS = SF-36 mental component summary score; MH = SF-36 mental health domain score; PRaCTICAL = a pragmatic randomized, controlled trial of intensive care follow-up programs in improving longer-term outcomes from critical illness; PTSD = post-traumatic stress disorder.

Data presented as mean (SD) unless otherwise specified.

\*Originally reported as APACHE III (mean, 86; SD, 25) and presented as APACHE II using standard conversion (44).

<sup>†</sup>We defined instrument-specific thresholds for substantial symptoms: ≥8 for the depression and anxiety subscales of HADS; ≥21 for DASS depression; ≥15 for DASS anxiety; ≥27 for DTS; ≥20 for IES; and ≥1.6 for IES-R.

<sup>‡</sup>Measured using the HADS.

<sup>§</sup>Measured using the DASS.

<sup>||</sup>Measured using the IES-R.

<sup>¶</sup>Measured using the IES.

<sup>\*\*</sup>Measured using the DTS.

studies. Participants were 53% men, with a mean age of 51 years, with some variability in mean age, severity of illness, ventilation duration, and length of stay across studies (Table 1). The number of patients completing each psychological instrument at the 6-month assessment were: 1,068 and 159 using HADS and DASS, respectively, as measures for depression and anxiety symptoms, and 763 using IES-R, 219 using DTS, and 159 using IES as PTSD measures.

At the 6-month assessment, the percentage of survivors who met the threshold for substantial symptoms, using the various psychological instruments in each study, ranged from 9 to 36% for depression, 11 to 42% for anxiety, and 19 to 42% for PTSD (Table 1). In pooled analyses among studies using the same instrument at the 6-month assessment, the proportion of patients with substantial symptoms for depression, anxiety, and PTSD were 32, 40, and 23% using HADS-depression ( $n = 335$ ), HADS-anxiety ( $n = 420$ ), and IES-R ( $n = 178$ ), respectively, with little change when analyses were replicated at 12 months (31%,  $n = 274$ ; 39%,  $n = 337$ ; 22%,  $n = 149$ , respectively).

A histogram of the MH and MCS scores across all studies is included in Figure E1 in the online supplement.

### Correlation between SF-36 and Psychological Instruments

In the pooled analysis, at the 6-month assessment, there were strong correlations of both the SF-36 MH and MCS scores with HAD depression and anxiety symptoms ( $r = -0.72$  to  $-0.79$ ) and strong correlations with IES-R PTSD symptoms ( $r = -0.60$  to  $-0.62$ ) (Table E1, Figure E2). Among the psychological instruments that were only measured in a single study, the DASS anxiety and IES scales had moderate correlations (DASS,  $r = -0.57$  and  $-0.53$ ; IES  $r = -0.54$  and  $-0.50$ ), whereas the DASS depression and DTS had strong correlations with the MH and MCS ( $-0.75$  and  $-0.73$ ,  $-0.75$  and  $-0.71$ , respectively).

Results were similar at the 12-month assessment (Table 2). The sensitivity analysis removing the CESAR dataset from pooled data did not change the point estimates when compared with the MH ( $r = -0.74$ ,  $-0.79$ , and  $-0.60$  for depression, anxiety, and PTSD, respectively) or MCS ( $r = -0.72$ ,  $-0.75$ , and  $-0.62$  for depression, anxiety, and PTSD, respectively) scores.

### Performance of SF-36 in Identifying Substantial Psychological Symptoms

At the 6-month assessment, the AUROC was 0.92 for DASS depression, 0.88 for HADS depression, 0.90 for DASS anxiety, 0.91 for HADS anxiety, 0.77 for IES, 0.84 for IES-R, and 0.85 for DTS (Figure 1). Similar results were observed at 12 months, with an AUROC of 0.89 for HADS depression, 0.90 for HADS anxiety, 0.85 for IES-R, and 0.90 for DTS.

At 6 months, using a threshold of 50 (i.e., the normalized mean score), the SF-36 MH domain score generally had a high sensitivity for the depression and anxiety subscales (range, 90–94%), with a lower sensitivity for the PTSD IES-R and DTS scales (78–92%). Specificity was lower than sensitivity across all scales (49–72% for depression and anxiety and 57–72% for PTSD). Positive predictive values were similar across symptoms of depression (16 and 55% for DASS and HADS-D, respectively), anxiety (20 and 68% for DASS and HADS-A, respectively), and PTSD (40, 44, and 67% for IES-R, IES, and DTS, respectively).

Some psychological instruments produced smaller ranges of positive predictive values that were lower

**Table 2.** Spearman correlation of Medical Outcomes Study Short Form-36 mental health scale and summary scores with psychological instruments

Study	Depression				Anxiety				PTSD				
	Instrument	No.*	Mean (SD)	MH <sup>†</sup>	MCS	Instrument	Mean (SD)	MH <sup>†</sup>	MCS	Instrument	Mean (SD)	MH <sup>†</sup>	MCS
6 mo													
Pooled <sup>‡</sup>	HADS-D	≥763	5.7 (5)	-0.74 <sup>§</sup>	-0.72 <sup>§</sup>	HADS-A	6.7 (5)	-0.79 <sup>§</sup>	-0.75 <sup>§</sup>	IES-R	1.0 (0.9)	-0.60 <sup>§</sup>	-0.62 <sup>§</sup>
ALTOS	HADS-D	≥603	6.1 (5)	-0.77	-0.74	HADS-A	7.1 (5)	-0.79	-0.77	IES-R	1.0 (0.9)	-0.63	-0.65
ICAP	HADS-D	≥160	5.4 (4)	-0.65	-0.64	HADS-A	5.6 (5)	-0.75	-0.70	IES-R	0.8 (0.8)	-0.46	-0.49
CESAR	HADS-D	74	4.9 (4)	-0.62	—	HADS-A	4.9 (4)	-0.66	—	—	—	—	—
PRaCTICAL	HADS-D	≥212	5.3 (4)	-0.72	-0.72	HADS-A	6.5 (5)	-0.81	-0.73	DTS	31.2 (31)	-0.75	-0.71
Elliot	DASS	159	7.8 (8)	-0.75	-0.73	DASS	6.3 (7)	-0.57	-0.53	IES	14.4 (16)	-0.54	-0.50
12 mo													
Pooled <sup>‡</sup>	HADS-D	≥708	5.6 (5)	-0.75 <sup>§</sup>	-0.75 <sup>§</sup>	HADS-A	6.7 (5)	-0.80 <sup>§</sup>	-0.7 <sup>§</sup>	IES-R	1.0 (0.9)	-0.61 <sup>§</sup>	-0.62 <sup>§</sup>
ALTOS	HADS-D	≥567	5.9 (5)	-0.76	-0.75	HADS-A	7.0 (5)	-0.80	-0.78	IES-R	1.0 (0.9)	-0.63	-0.65
ICAP	HADS-D	≥141	5.2 (4)	-0.72	-0.70	HADS-A	6.4 (5)	-0.78	-0.76	IES-R	0.9 (0.9)	-0.52	-0.51
PRaCTICAL	HADS-D	≥187	4.8 (4)	-0.74	-0.78	HADS-A	6.0 (5)	-0.80	-0.77	DTS	26.7 (28)	-0.77	-0.75

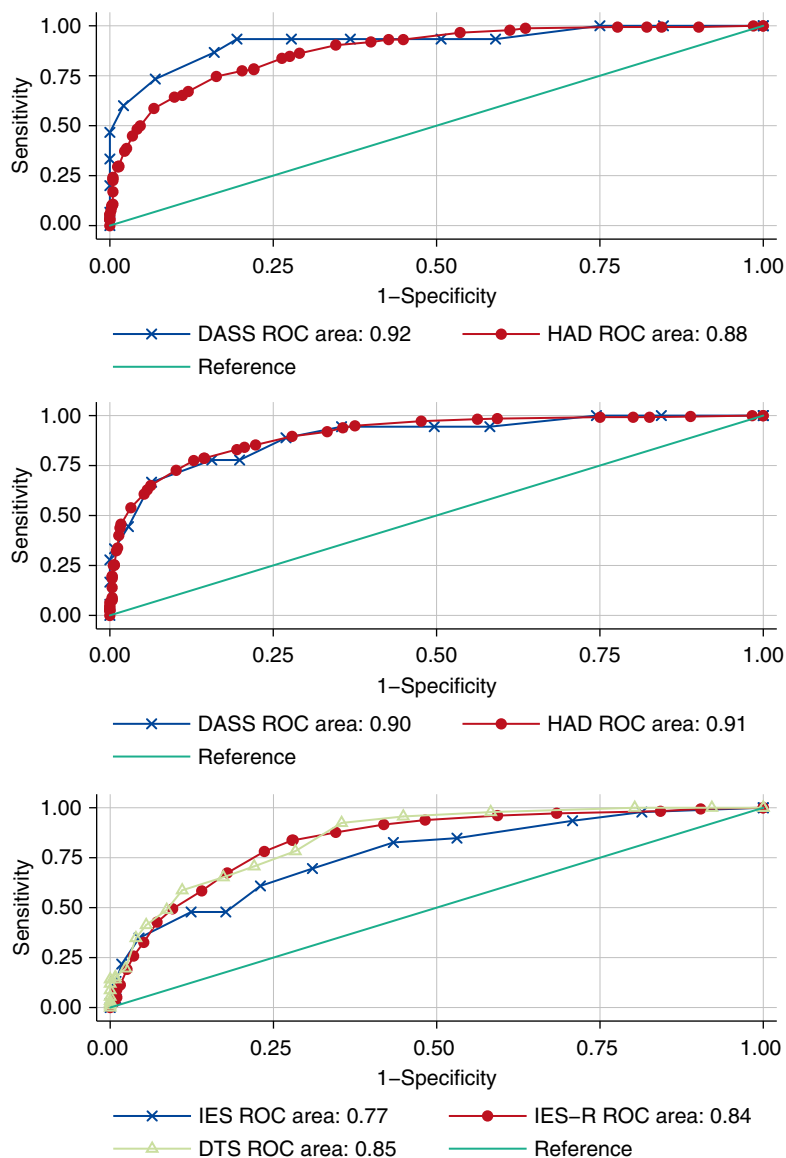
*Definition of abbreviations:* ALTOS = ARDSNet Long Term Outcomes study; CESAR = conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure; DASS = Depression Anxiety Stress Scales; DTS = Davidson Trauma Scale; HADS = Hospital Anxiety and Depression Scale; ICAP = Improving Care of Acute Lung Injury Patients study; IES = Impact of Event Scale; IES-R = Impact of Event Scale-Revised; MCS = SF-36 mental component summary score; MH = SF-36 mental health domain score; PRaCTICAL = a pragmatic randomized, controlled trial of intensive care follow-up programs in improving longer-term outcomes from critical illness; PTSD = post-traumatic stress disorder.

\*Sample size may vary across analyses because not all participants have all measures; hence, the minimum sample size across all analyses is presented in each row.

<sup>†</sup>Presented as a normalized score (mean, 50; SD, 10).

<sup>‡</sup>Pooled data include ALTOS, ICAP, CESAR, and PRaCTICAL for depression and anxiety, and ALTOS and ICAP for PTSD.

<sup>§</sup> $P$  value  $\leq 0.01$ .



**Figure 1.** Receiver operating characteristic (ROC) curves of Medical Outcomes Study Short Form (SF)-36 mental health domain score versus substantial symptoms of depression (*top*), anxiety (*middle*), and post-traumatic stress disorder (*bottom*) at the 6-month assessment. DASS = Depression Anxiety Stress Scales; DTS = Davidson Trauma Scale; HAD = Hospital Anxiety and Depression; IES = Impact of Event Scale; IES-R = Impact of Event Scale-Revised.

(e.g., 11–36% for DASS depression) or higher (e.g., 45–83% for HADS anxiety) than the range across all instruments, but this result may be due to different prevalence of psychological conditions in these individual studies. Negative predictive values were high across all instruments (82–99%). If a lower cut-point on the SF-36 were used (e.g., a cut-point of 40), positive predictive values increase substantially without a parallel decrease of negative predictive values (Table E1). These results were similar when replicated to evaluate the

SF-36 MH score at the 12-month assessment and the SF-36 MCS (rather than MH score) at the 6- and 12-month assessments (*see online supplement*).

#### Factor Analysis

Factor loadings were similar at 6- and 12-month assessments, and they were similar to those in the U.S. general population (36) (Table E2). As seen in the U.S. general population (36), the MH loaded strongly onto the MCS and weakly onto the physical health

component summary score. This finding suggests that the SF-36 measures the constructs of physical and mental health-related quality of life similarly in survivors of acute respiratory failure as in the general population in the United States.

## Discussion

Our study presents results from a tri-national analysis of five studies of survivors of acute respiratory failure over their first year of follow up. When compared with evaluations with commonly used psychological self-report instruments, the SF-36 MH and MCS scores both demonstrated strong correlations with depression and anxiety symptoms and moderate correlations with PTSD symptoms. Across different psychological instruments measuring the same psychological phenomenon, performance of the SF-36 MH and MCS was similar, although associations were reduced with the PTSD scales.

The factor analysis, along with prior validation of the SF-36 (9, 10), suggests that the structure of the SF-36 is similar in survivors of acute respiratory failure to a representative general population sample (36). Overall, our results support the use of the SF-36 in survivors of acute respiratory failure for researchers interested in general mental health status. However, for those specifically interested in depression, anxiety, and PTSD symptoms, specific psychological scales validated in survivors of acute respiratory failure should be used.

The SF-36 had high sensitivity and negative predictive values at the cut-point of 50 and lower specificity and positive predictive values; therefore, using a cut-point of 50, researchers would identify more people as screening positive for psychological symptoms than if survivors were screened using specific psychological instruments.

Researchers should choose a cut-point on the SF-36 on the basis of their objective. For example, using a cut-point of 40, 83% of patients with a score less than or equal to 40 would screen positive for anxiety on the HADS-A. Conversely, for those interested in ruling out individuals with good general mental health, a higher cut-point could be used; for example, a cut-point

of 60 would result in 98% negative predictive value for depression and anxiety when compared with the HADS-A. Of note, because negative predictive values were generally high across all instruments and cut-points, the best use of the SF-36, in a research setting, is to help rule out patients with mental health disorders in studies.

The SF-36 MH and MCS, which were designed to measure general mental health status, include four major mental health dimensions: anxiety, depression, loss of behavioral/emotional control, and psychological well-being (8). The inclusion of anxiety and depression in the SF-36 reflect overlap with the specific psychological instruments that measure depression and anxiety used in this study. Consequently, strong correlations of the SF-36 MH and MCS scores with a number of anxiety- and depression-specific instruments at both 6 and 12 months are not surprising. Previous work also found that the SF-36 mental health domain strongly correlates with the HADS anxiety subscale at 3 months in survivors of acute respiratory failure (40). Together, these results provide evidence of stability of the relationship between the SF-36 and anxiety- and depression-specific instruments over the first 12 months of patient follow up.

In contrast to the above results, the construct of PTSD encompasses intrusion, avoidance, and alterations in arousal and reactivity symptoms (41), which are less comprehensively evaluated in the SF-36, likely explaining the generally moderate (rather than strong) relationship of the SF-36 with the PTSD-specific instruments and values for AUROC curves that were fair when compared with the IES (0.77) and good with the IES-R and the DTS (0.84 and 0.85, respectively). To have more specific evaluation of PTSD symptoms, interested researchers should use a PTSD-specific instrument. Previous research has validated the IES-R and the Post-Traumatic Stress Syndrome 10-Questions Inventory (vs. semistructured diagnostic interviews) (33, 42) and the UK-Post-Traumatic Stress Syndrome 14-Questions Inventory (vs. the Post-traumatic Stress Diagnostic Scale) (43) for PTSD in this population.

Our results also demonstrated that the MH and MCS scores had similar relationships across different psychological

instruments and at both the 6- and 12-month assessment points. In the correlation analyses, the MH and MCS produced results that were not meaningfully different. For example, the correlations between the HADS depression subscale versus the SF-36 MH and MCS scores, at the 12-month assessment, were both  $-0.75$ . The largest difference in association was found in the PRaCTICAL study's 6-month HADS anxiety assessment with a correlation of  $-0.73$  with MH score and  $-0.81$  with MCS score. Furthermore, the sensitivity and specificity analysis demonstrated nearly identical results for the MH and MCS. This finding supports the use of either the MH or the MCS in the research setting. Given that the MH requires fewer questions than the MCS (11), using only the MH questions may reduce respondent burden.

### Strengths and Limitations

Our study has several strengths. The study includes data from five studies across three continents and two time points, increasing the external generalizability of the results. In addition, we repeated the analysis for both the MH and MCS scales. Across these different studies and time points, our results were similar, demonstrating stability of our findings.

This study also has limitations. First, we compared the SF-36 to self-report psychological screening instruments for depression, anxiety, and PTSD. This analysis is not intended to suggest that the SF-36 could be used to make depression-, anxiety-, or PTSD-related diagnoses, nor should it replace measures of depression, anxiety, and PTSD in every study. Furthermore, the results of this study should not be used to inform diagnostic procedures in the clinical setting. Instead, we intended this analysis to provide researchers new information to consider when selecting measurement instruments to assess general mental health in acute respiratory failure survivorship research.

Second, the SF-36 cannot differentiate the constructs of depression and anxiety, meaning that it cannot be used to differentiate specific psychological conditions; rather, it is a more general measure of mental health, as initially intended with its design. Third, we only used U.S. population norms in the factor analysis

of the SF-36. However, previous work derived similar factors from the SF-36 in the United Kingdom (37, 38) and Australia (39); hence, these results should be generalizable across this tri-national study.

Fourth, except for the IES-R (33), and some limited analyses with the HAD (40), the psychological screening instruments used in this study have not been validated in survivors of acute respiratory failure. The uncertain association of the screening instruments with clinical diagnoses may limit inferences from the analysis. However, the instruments are commonly accepted and in widespread use and are of value in understanding the magnitude of relevant psychological symptoms.

Finally, the generalizability of this study is limited. The study was only conducted in survivors of acute respiratory failure, there are fewer participants at the 12-month time point, and not all psychological measures were administered in all patients, with some measures having relatively smaller sample sizes.

### Conclusions

In conclusion, the SF-36 MH and MCS scores offer strong measures of mental health symptoms frequently experienced by survivors of acute respiratory failure in their first year of recovery on the basis of the correlation analysis and the AUROC curves across multiple studies, acute respiratory failure patient populations, and screening instruments. Studies interested in measuring the general mental health status of survivors of acute respiratory failure may consider using either the MCS or MH, although the MH domain score may be preferable given the fewer questions required. Researchers with interest in specific psychological conditions should consider specific psychological scales validated in survivors of acute respiratory failure, rather than the SF-36 alone. ■

**Author disclosures** are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).

**Acknowledgment:** The authors thank all the patients and their proxies who participated in the original studies. They also thank the investigators and research staff from the ARDS Network, ALTOS, ICAP, CESAR, PRaCTICAL, and the trial by Elliott and colleagues for their important contributions. They also thank Minxuan Huang for his help with data management.

## References

- Davydow DS, Katon WJ, Zatzick DF. Psychiatric morbidity and functional impairments in survivors of burns, traumatic injuries, and ICU stays for other critical illnesses: a review of the literature. *Int Rev Psychiatry* 2009;21:531–538.
- Davydow DS, Gifford JM, Desai SV, Bienvenu OJ, Needham DM. Depression in general intensive care unit survivors: a systematic review. *Intensive Care Med* 2009;35:796–809.
- Parker AM, Srichaonchai T, Raparla S, Schneck KW, Bienvenu OJ, Needham DM. Posttraumatic stress disorder in critical illness survivors: a metaanalysis. *Crit Care Med* 2015;43:1121–1129.
- Spragg RG, Bernard GR, Checkley W, Curtis JR, Gajic O, Guyatt G, Hall J, Israel E, Jain M, Needham DM, et al. Beyond mortality: future clinical research in acute lung injury. *Am J Respir Crit Care Med* 2010;181:1121–1127.
- Angus DC, Carlet J; 2002 Brussels Roundtable Participants. Surviving intensive care: a report from the 2002 Brussels Roundtable. *Intensive Care Med* 2003;29:368–377.
- Dowdy DW, Eid MP, Sedrakyan A, Mendez-Tellez PA, Pronovost PJ, Herridge MS, Needham DM. Quality of life in adult survivors of critical illness: a systematic review of the literature. *Intensive Care Med* 2005;31:611–620.
- Dowdy DW, Eid MP, Dennison CR, Mendez-Tellez PA, Herridge MS, Guallar E, Pronovost PJ, Needham DM. Quality of life after acute respiratory distress syndrome: a meta-analysis. *Intensive Care Med* 2006;32:1115–1124.
- Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–483.
- Chrispin PS, Scotton H, Rogers J, Lloyd D, Ridley SA. Short Form 36 in the intensive care unit: assessment of acceptability, reliability and validity of the questionnaire. *Anaesthesia* 1997;52:15–23.
- Heyland DK, Hopman W, Coo H, Tranmer J, McColl MA. Long-term health-related quality of life in survivors of sepsis. Short Form 36: a valid and reliable measure of health-related quality of life. *Crit Care Med* 2000;28:3599–3605.
- Ware JE Jr. SF-36 health survey update. *Spine* 2000;25:3130–3139.
- Blackwood B, Marshall J, Rose L. Progress on core outcome sets for critical care research. *Curr Opin Crit Care* 2015;21:439–444.
- Needham DM. Understanding and improving clinical trial outcome measures in acute respiratory failure. *Am J Respir Crit Care Med* 2014;189:875–877.
- Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:132.
- Elliott D, McKinley S, Alison J, Aitken LM, King M, Leslie GD, Kenny P, Taylor P, Foley R, Burmeister E. Health-related quality of life and physical recovery after a critical illness: a multi-centre randomised controlled trial of a home-based physical rehabilitation program. *Crit Care* 2011;15:R142.
- Needham DM, Dinglas VD, Bienvenu OJ, Colantuoni E, Wozniak AW, Rice TW, Hopkins RO; NIH NHLBI ARDS Network. One year outcomes in patients with acute lung injury randomised to initial trophic or full enteral feeding: prospective follow-up of EDEN randomised trial. *BMJ* 2013;346:f1532.
- National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Matthay MA, Brower RG, Carson S, Douglas IS, Eisner M, Hite D, Holets S, Kallet RH, Liu KD, et al. Randomized, placebo-controlled clinical trial of an aerosolized  $\beta_2$ -agonist for treatment of acute lung injury. *Am J Respir Crit Care Med* 2011;184:561–568.
- Rice TW, Wheeler AP, Thompson BT, deBoisblanc BP, Steingrub J, Rock P; NIH NHLBI Acute Respiratory Distress Syndrome Network of Investigators. Enteral omega-3 fatty acid, gamma-linolenic acid, and antioxidant supplementation in acute lung injury. *JAMA* 2011;306:1574–1581.
- National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Rice TW, Wheeler AP, Thompson BT, Steingrub J, Hite RD, Moss M, Morris A, Dong N, Rock P. Initial trophic vs full enteral feeding in patients with acute lung injury: the EDEN randomized trial. *JAMA* 2012;307:795–803.
- Fan E, Dowdy DW, Colantuoni E, Mendez-Tellez PA, Sevransky JE, Shanholtz C, Himmelfarb CR, Desai SV, Ciesla N, Herridge MS, et al. Physical complications in acute lung injury survivors: a two-year longitudinal prospective study. *Crit Care Med* 2014;42:849–859.
- Peek GJ, Mugford M, Tiruvoipati R, Wilson A, Allen E, Thalanany MM, Hibbert CL, Truesdale A, Clemens F, Cooper N, et al.; CESAR trial collaboration. Efficacy and economic assessment of conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure (CESAR): a multicentre randomised controlled trial. *Lancet* 2009;374:1351–1363.
- Cuthbertson BH, Rattray J, Campbell MK, Gager M, Roughton S, Smith A, Hull A, Breeman S, Norrie J, Jenkinson D, et al. The PRaCTICaL study of nurse led, intensive care follow-up programmes for improving long term outcomes from critical illness: a pragmatic randomised controlled trial. *BMJ* 2009;339:b3723.
- Cuthbertson BH, Hull A, Strachan M, Scott J. Post-traumatic stress disorder after critical illness requiring general intensive care. *Intensive Care Med* 2004;30:450–455.
- Hawthorne G, Osborne RH, Taylor A, Sansoni J. The SF36 Version 2: critical analyses of population weights, scoring algorithms and population norms. *Qual Life Res* 2007;16:661–673.
- Maruish ME. User's manual for the SF-36v2 Health Survey. Lincoln, RI: Quality Metric Incorporated; 2011.
- Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361–370.
- Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav Res Ther* 1995;33:335–343.
- McKinley S, Aitken LM, Alison JA, King M, Leslie G, Burmeister E, Elliott D. Sleep and other factors associated with mental health and psychological distress after intensive care for critical illness. *Intensive Care Med* 2012;38:627–633.
- Davidson JR, Book SW, Colket JT, Tupler LA, Roth S, David D, Hertzberg M, Mellman T, Beckham JC, Smith RD, et al. Assessment of a new self-rating scale for post-traumatic stress disorder. *Psychol Med* 1997;27:153–160.
- Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. *Psychosom Med* 1979;41:209–218.
- Weiss DS. The Impact of Events Scale-Revised. In: Wilson JP, Keane TM, editors. Assessing psychological trauma and PTSD. New York: Guilford; 1997.
- Myhren H, Ekeberg O, Tøien K, Karlsson S, Stokland O. Posttraumatic stress, anxiety and depression symptoms in patients during the first year post intensive care unit discharge. *Crit Care* 2010;14:R14.
- Bienvenu OJ, Williams JB, Yang A, Hopkins RO, Needham DM. Posttraumatic stress disorder in survivors of acute lung injury: evaluating the impact of Event Scale-Revised. *Chest* 2013;144:24–31.
- TheBMJ. 11. Correlation and regression. [accessed 2016 Jun 22]. Available at: <http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression>
- Harshman RA, Lundy ME. PARAFAC: parallel factor analysis. *Comput Stat Data Anal* 1994;18:15–37.
- Maglinde GA, Hays RD, Kaplan RM. US general population norms for telephone administration of the SF-36v2. *J Clin Epidemiol* 2012;65:497–502.
- Jenkinson C. Comparison of UK and US methods for weighting and scoring the SF-36 summary measures. *J Public Health Med* 1999;21:372–376.
- Ware JE Jr, Kosinski M, Gandek B, Aaronson NK, Apolone G, Bech P, Brazier J, Bullinger M, Kaasa S, Leplège A, et al. The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. International Quality of Life Assessment. *J Clin Epidemiol* 1998;51:1159–1165.

- 39 Sanson-Fisher RW, Perkins JJ. Adaptation and validation of the SF-36 Health Survey for use in Australia. *J Clin Epidemiol* 1998;51:961–967.
- 40 Jutte JE, Needham DM, Pfoh ER, Bienvenu OJ. Psychometric evaluation of the Hospital Anxiety and Depression Scale 3 months after acute lung injury. *J Crit Care* 2015;30:793–798.
- 41 Friedman MJ. Finalizing PTSD in DSM-5: getting here from there and where to go next. *J Trauma Stress* 2013;26:548–556.
- 42 Stoll C, Kapfhammer HP, Rothenhäusler HB, Haller M, Briegel J, Schmidt M, Krauseneck T, Durst K, Schelling G. Sensitivity and specificity of a screening test to document traumatic experiences and to diagnose post-traumatic stress disorder in ARDS patients after intensive care treatment. *Intensive Care Med* 1999;25:697–704.
- 43 Twigg E, Humphris G, Jones C, Bramwell R, Griffiths RD. Use of a screening questionnaire for post-traumatic stress disorder (PTSD) on a sample of UK ICU patients. *Acta Anaesthesiol Scand* 2008;52:202–208.
- 44 Schneider AG, Lipcsey M, Bailey M, Pilcher DV, Bellomo R. Simple translational equations to compare illness severity scores in intensive care trials. *J Crit Care* 2013;28:885.e1–885.e8.